ODICAL ROOM
ERAL LIBRARY
Y. OF MICH.
MAY 9 1923

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS

E. J. ASHBAUGH W. W. CHARTERS GEORGE D. STRAYER
GUY M. WHIPPLE S. A. COURTIS LEWIS M. TERMAN
FRANK N. FREEMAN

VOL. VII APRIL, 1923 No. 4

CONTENTS

	Page
RELATION OF GENERAL INTELLIGENCE TO THE PERSISTENCE OF EDUCATIONAL AND VOCATIONAL PLANS OF HIGH-SCHOOL PUPILS — William Martin Proctor and Helen Ward	277
FINANCING SCHOOLS UNDER THE DISTRICT SYSTEM — Fletcher Harper Swift	289
THE ANALYSIS OF TEACHING FUNCTIONS — Frederick L. Whitney	297
A SURVEY OF THE TEST MOVEMENT IN HISTORY — Paul Tyler Kepner	309
ANOTHER CRITICISM OF TESTS REQUIRING ALTERNATIVE RESPONSES — C. W. Odell	326
A MENTAL SURVEY OF 59 NORMAL-SCHOOL STUDENTS: SOME CORRELATIONS AND CRITICISMS — Lawrence A. Averill	331
TRANSMUTATION OF SCORES BETWEEN BINET TESTS AND GROUP TESTS — W. T. Root	338
EDITORIALS	344
REVIEWS AND ABSTRACTS	350
NEWS ITEMS AND COMMUNICATIONS	359
EDUCATIONAL RESEARCH ASSOCIATION	367

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1923, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, B. R. Buckingham
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ohio State University, Columbus, Ohio.*

W. W. CHARTERS, *Carnegie Inst. Tech., Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

FRANK N. FREEMAN, *Ex-Officio, Secretary Educational Research Association.*

GEORGE D. STRAYER, *Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California*

GUY M. WHIPPLE, *University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska, Lincoln, Nebr.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota, Minneapolis.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin, Madison.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean of University Administration, University of Kansas, Lawrence, Kans.*

ROBT. H. LANE, *Assistant Superintendent of Schools Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.*

LEON O. SMITH, *Assistant Superintendent (in charge of Educational Research) Public Schools, Omaha, Nebraska.*

W. W. THEISEN, *Director, Division of Reference and Research, Public Schools, Cleveland, Ohio.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VII APRIL, 1923 Number 4

RELATION OF GENERAL INTELLIGENCE TO THE PERSISTENCE OF EDUCATIONAL AND VOCATIONAL PLANS OF HIGH-SCHOOL PUPILS

WILLIAM MARTIN PROCTOR AND HELEN WARD
Stanford University

The question is frequently raised as to the value of data concerning the present ambitions of high-school pupils. The claim is made that the answers to questions regarding vocational and educational plans are unreliable either because the pupil has no definite ideas regarding his life work or because he does not take the question seriously. It is therefore of genuine interest to all who attempt the advisement of youth to know whether and to what extent the vocational ambitions and educational plans of boys and girls in the high school represent significant and abiding life interests.

The writers have just completed the second check-up on a group of high-school pupils whose vocational ambitions, educational plans, school success, and general intelligence scores were secured in the year 1917-18. The first check-up was made two years after the original data were secured, and the results were reported in a monograph which dealt primarily with the possibilities of utilizing mental tests as an aid in the educational and vocational guidance of high-school pupils.[1] The present check-up represents the situation as it exists in 1921-22, four years after the mental tests were given and the questions regarding educational and vocational plans were answered.

[1] Proctor, W. M., *Psychological tests and guidance of high-school pupils.* (Journal of Educational Research Monographs, No. 1, June, 1921.) Bloomington, Illinois: Public School Publishing Company, 1921. 70 pp.

Only those pupils who attended the Palo Alto, San Mateo, Redwood City, Mountain View, and Santa Clara high schools are considered in the present study. It was found to be difficult to secure information concerning cases from the other high schools which were included in the original investigation. An additional item in favor of confining the fourth-year check-up to the five high schools we have named, was the fact that there had not been a change in administrative head in any of these high schools during the four-year period.

In these high schools 771 pupils had originally filled out the questionnaires in full and had taken the mental tests. Some 88 of these could not be located, leaving 683 concerning whom reliable information could be had. The 683 cases were distributed as follows: (a) in occupations, 272; (b) in educational institutions beyond the high school, 290; (c) still in the high school, 33; (d) married, 47; (e) just graduated, plans indefinite, 27; (f) prevented from carrying out plans by ill health, 10; (g) died, 4.

The mental age and I.Q. values hereafter referred to are those found in the original examination. No new tests were given. The Army Alpha scale was the group test employed, and the mental-age equivalents were derived from the Kohs-Proctor mental-age norms for the Alpha scale.[2]

Persistence of Vocational Plans of Cases Now in Occupations

Table I gives the distribution of the 272 cases found to be engaged in occupations, according to their intelligence levels and also as to the relation of the present occupation to the vocational preference expressed four years ago. Sex differences are indicated under the captions "B" and "G." Wherever in this and in subsequent tables the rank of vocations is referred to it is to be understood that occupations are ranked from I to V as follows: Rank I, higher professional and executive positions;

[2] Proctor, W. M., "Psychological tests as a means of measuring the probable school success of high-school pupils." *Journal of Educational Research.* 1:258-70 (261). April, 1920.

Rank II, semi-professional, managerial, and higher commercial positions; Rank III, general clerical and commercial positions, skilled labor, etc.; Rank IV, semi-skilled labor; Rank V, unskilled labor.

TABLE I. RELATION OF OCCUPATION IN 1921-22 TO PREFERENCE IN 1917-18

	General Intelligence—I. Q. (Alpha)						Totals		
	80–99		100–119		120 or over				
	B	G	B	G	B	G	B	G	Both
Vocation same as preference stated in 1917-18	3	28	14	34	3	5	20	67	87
Different vocation but of same rank as original preference	1	7	4	5	1	1	6	13	19
Higher rank	0	0	0	0	0	2	0	2	2
Lower rank	39	31	46	31	15	2	100	64	164
Totals	43	66	64	70	19	10	126	146	272

When the data of Table I are summarized and expressed in the form of percents, the meaning of the table becomes more evident. Forty percent of these former high-school pupils are engaged in occupations of the same or higher rank with reference to their original vocational preference. Correspondingly, 60 percent are in occupations of lower rank than that of the preference expressed in 1917-18. The sex differences are quite marked, especially when taken in connection with the intelligence levels. For example, among the boys, 9 percent of those with I.Q.'s below 100 and 30 percent of those with I.Q.'s above 100 are engaged in occupations equal to or above their preference. On the other hand, among the girls, those above average intelligence have been no more successful than those of less than average intelligence in attaining their objectives. The percents are 51 for the former and 53 for the latter. In other words, intelligence appears to have played a greater part in the case of the boys than it did in the case of the girls.

The sex differences brought out in the preceding paragraph are accounted for in part by the fact that in the original survey the vocational ambitions of the boys ranked higher than those of the girls. This point will be further discussed in connection with Table II.

As has already been shown in connection with Table I the percent of girls with I.Q.'s below 100, who were in occupations of their choice or in occupations of higher rank was 53, while the percent of boys in the same situation was only 9. In Table III, however, it appears that 56 percent of the boys and 38 percent of the girls with I.Q.'s from 80-99 had vocational prefer-

TABLE II. (a) THE ORIGINAL VOCATIONAL PREFERENCES AS EXPRESSED IN 1917-18

Rank of Occupations Preferred	General Intelligence—I. Q.						Totals		
	80–99		100–119		120 or over				
	B	G	B	G	B	G	B	G	Both
I	13	2	22	4	13	0	48	6	54
II	11	23	20	35	3	5	34	63	97
III	19	41	22	31	2	5	43	77	120
IV	0	0	0	0	1	0	1	0	1
V	0	0	0	0	0	0	0	0	0
Totals.....	43	66	64	70	19	10	126	146	272

(b) OCCUPATIONS ENGAGED IN FOUR YEARS LATER (1921-22)

Rank of Present Occupation	General Intelligence—I. Q.						Totals		
	80–99		100–119		120 or over				
	B	G	B	G	B	G	B	G	Both
I	0	0	0	0	0	0	0	0	0
II	2	3	4	5	3	3	9	11	20
III	14	50	22	50	4	7	40	107	147
IV	15	13	30	15	11	0	56	28	84
V	12	0	8	0	1	0	21	0	21
Totals.....	43	66	64	70	19	10	126	146	272

TABLE III. PERCENTAGE SUMMARY OF TABLE II. (a) ORIGINAL VOCATIONAL PREFERENCES EXPRESSED IN 1917-18

Rank of Occupations Preferred	General Intelligence—I. Q.						Totals		
	80–99			100 or over					
	B	G	Both	B	G	Both	B	G	Both
I	30.5	3.0	16.0	42.1	5.0	24.0	38.0	4.0	20.0
II	25.5	35.0	31.0	27.7	50.0	39.0	27.2	43.0	35.6
III	44.0	62.0	53.0	29.0	45.0	37.0	34.0	53.0	44.0
IV	0	0	0	1.2	0	0	0.8	0	0.4
V	0	0	0	0	0	0	0	0	0

(b) OCCUPATIONS ENGAGED IN WHEN CHECKED UP IN 1921-22

Rank of Occupation now Engaged in	General Intelligence—I. Q.						Totals		
	80–99			100 or over					
	B	G	Both	B	G	Both	B	G	Both
I	0	0	0	0	0	0	0	0	0
II	4.5	4.5	4.5	8.4	10.0	9.0	7.0	8.0	7.0
III	32.5	75.5	58.8	31.4	71.0	51.0	32.0	73.0	54.0
IV	35.0	20.0	25.7	49.2	19.0	34.0	44.0	19.0	31.0
V	28.0	0	11.0	11.0	0	6.0	17.0	0	8.0

ences that ranked above III, in 1917-18, while in 1921-22 63 percent of the boys and 20 percent of the girls in that intelligence group were in occupations ranking IV and V; in 1917-18 the median vocational preference of the boys was Rank II; of the girls Rank III. In 1921-22, the median occupation of the boys was Rank IV; of the girls Rank III.

When we consider the group with I.Q.'s of 100 or over, we note that whereas about 70 percent of the boys and 55 percent of the girls had ambitions ranking I and II, none of either sex was found in occupations ranking I, and only 9 percent in occupations ranking II. In 1917-18, none of the girls and only 1.2 percent of the boys had ambitions ranking IV or V, but in 1921-22, we find about half the boys in occupations ranking IV, and 11 percent in occupations ranking V. We find none of the

girls of this I.Q. group in occupations ranking V, but 19 percent of them are found in occupations ranking IV.

The influence of intelligence on persistence in vocational plans is apparently discoverable, but at this stage of the investigation it has not yet become pronounced. Our data permit a few pertinent statements. According to section (b) of Table III, 9 percent of those of better than average intelligence are in occupations ranking II, while only 4.5 percent of those of less than average intelligence are in occupations of that rank. Again, 28 percent of the boys with below average intelligence are in occupations ranking V, while only 11 percent of the boys with above average intelligence are in this lowest class of occupations. In the case of the boys, it has already been stated in connection with Table I that those of better than average intelligence showed greater persistence in the occupations of their original choice. The figures were 30 percent for those with I.Q.'s of 100 or higher and 9 percent for those with I.Q.'s below 100.

A check-up, made four years after the original survey, could hardly be expected to reveal pronounced correlations between intelligence and persistence in vocational plans, but it is significant that there seems to be a tendency for those of below average ability to gravitate toward occupations ranking IV and V. A similar check-up, made four years hence, will probably reveal a much more pronounced tendency for those of above average ability to be found in the higher ranking occupations than is now the case. With reference especially to boys, it is well known that they have to start at the bottom in a given occupation or that they may even have to take some temporary job which has no relation to their vocational preference.

The data here presented at least lends support to the supposition that those of better than average ability will tend gradually to approximate their vocational ambitions, while those with below average ability will be more apt to find their ultimate occupations at levels below their high-school plans for a life career. Two concrete cases will illustrate this point. A certain boy whose I.Q. was 92 expressed vocational preference, "Doctor of

Medicine." He is now driving a delivery wagon. Another whose ambition was that of "Electrical Engineer," had an I.Q. of 89. He left the high school because of failure in his work, and is now working at odd jobs as a day laborer.

Illustrating the tendency of those with above average ability to climb to higher occupational levels are the two cases in Table I, of girls who are following vocations of higher rank than their vocational preference. One, I. Q. 129, whose ambition was to be a stenographer, is now a teacher; the other, I.Q. 130, who was also ambitious to be a stenographer, started her work in that vocation, and is now a successful office manager.

Relation of Intelligence to Training before Entering Occupations

Table IV distributes the 272 persons engaged in occupations in relation to their intelligence and also in relation to the specific training which they had for the occupation which they entered.

TABLE IV. RELATION BETWEEN INTELLIGENCE AND OCCUPATIONAL TRAINING

Training for the Occupation before Entering it	Intelligence Ratings—I. Q.									Totals		
	80–99			100–119			120 or over					
	B	G	Both	B	G	Both	B	G	Both	B	G	Both
Some in High School or Later												
Number	9	24	33	9	18	27	8	3	11	26	45	71
Percent	21	36	30	14	26	20	42	30	38	21	31	26
None in High School or Later												
Number	34	42	76	55	52	107	11	7	18	100	101	201
Percent	79	64	70	86	74	80	58	70	62	79	69	74
Total	43	66	109	64	70	134	19	10	29	126	146	272

The girls evidently had more training for their occupations than the boys. This is expressed by the percent of 31 for the girls and 21 for the boys. Moreover, as has already been shown, a higher percent of girls than of boys tended to persist in their vocational ambitions. The character of the vocational ambitions of the girls may account for these facts. By far the greater number of girls were ambitious to fill clerical positions and

either took commercial subjects in the high school or attended a business college after graduation and before securing a position.

No unmistakable relation between intelligence and occupational training is apparent. On the one hand, more of the group of less than average intelligence had taken such training than was true of the average group. On the other hand, more of the superior group had occupational training than was the case with either of the other two groups. Differences, however, as revealed by the percents of Table IV are not large enough to be especially significant.

The real point of interest in this table is the fact that only 71 out of 272 of these former high-school pupils, or 26 percent of them, had any training either in the high school or after leaving the high school for the occupations in which they are now engaged. In other words, approximately three out of every four of them had not had any preparation whatever for their vocation. Either there was no vocational guidance worth mentioning in connection with these five high schools, or the provision for vocational instruction was very meager. The truth is that both counts of the indictment can be established beyond a reasonable doubt. These five high schools are under the shadow of Stanford University, and none of them is more than fifty miles from the University of California. Their location has caused them to stress college entrance to the neglect of vocational education. It is not, therefore, surprising that 74 percent of those who entered occupations after attending these high schools were compelled to do so without any specific preparation.

Relation between Intelligence and Persistence in Educational Plans

The reader will recall that 290 of the students on whom, after the four-year period, the data were secured were still attending some educational institution. In order to compare their plans as expressed in 1917-18 with their realization of them in 1921-22, an arbitrary rating of the educational institutions they were attending was applied. Thus a college or university was rated higher than a normal school, a normal school higher than

a business college, etc. Table V shows the relation between the intelligence of these students and the rank of the institutions they were attending 1921-22.

In contrast with the showing of Table I, where it was found that only 40 percent of those who entered occupations were employed in the occupation of their original choice, Table V indicates that 81 percent of those who continued their education beyond the high school are carrying out plans the same as, equal to, or superior to their original plans for an education. Also it appears from Table V that 89 percent of the boys and only 75 percent of the girls are following the same or higher educational plans, whereas of those in occupations only 26 percent of the boys and 56 percent of the girls were carrying out their original intentions.

The difference in persistence of educational plans between the boys and girls is due in a measure to the fact that most of the boys have ambitions leading to professional occupations that require a college education, while a considerable number of girls who originally planned to go to college, decided, after graduation from the high school, to teach and are now attending normal schools.

There is a more clearly marked relationship between intelli-

TABLE V. DISTRIBUTION OF 290 PUPILS ACCORDING TO INTELLIGENCE AND PERSISTENCE IN EDUCATIONAL PLANS

	General Intelligence—I. Q.									Totals		
	80–99			100–119			120 or over					
	B	G	Both	B	G	Both	B	G	Both	B	G	Both
Pupils Following Educational Plans the Same as, Equal to, or Superior to Original Plans												
Number	9	21	30	57	70	127	53	25	78	119	116	235
Percent	64	62	63	85	86	81	98	83	93	89	75	81
Pupils Following Educational Plans of Lower than Original Rank												
Number	5	13	18	10	21	31	1	5	6	16	39	55
Percent	36	38	37	15	14	19	2	17	7	11	25	19
Totals	14	34	48	67	91	158	54	30	84	135	155	290

gence and persistence in educational plans than there is between intelligence and persistence in vocational plans. There is a decided rise in the percent of boys and girls following the same or higher educational plans as the average of intelligence increases. In the lowest intelligence group, 63 percent; in the medium group, 81 percent; and in the highest group, 93 percent have persisted in the same or better plans for their education.

There appears to be a clear-cut difference in intelligence between the pupils now in occupations and those in educational institutions beyond the high school. This fact is shown in general by the median intelligence quotients of the two groups. For those in educational institutions it is 115, and for those in occupations it is 105. The condition is exhibited in greater detail in Table VI.

TABLE VI. INTELLIGENCE OF THOSE IN OCCUPATIONS AND IN EDUCATIONAL INSTITUTIONS

	General Intelligence—I.Q.					
	80–99		100–119		120 or over	
	Number	Percent	Number	Percent	Number	Percent
Group now in educational institutions ..	48	16	158	55	84	29
Group now in occupations.....	109	40	134	49	29	11

In this table the reader's attention is particularly called to the fact that among the pupils now in educational institutions only 16 percent have intelligence quotients below 100, while on the other hand, of the pupils now in occupations 40 percent have intelligence quotients below 100. At the high end of the range of intelligence the preponderance is reversed. In other words, a much greater proportion of those in educational institutions have intelligence quotients of 120 or over than is the case with those now in occupations.

Another distribution of the group in occupations—a distribution which is not here reproduced in tabular form—brought out the general fact that those who left the high school to go to

work before graduation had a median intelligence quotient of 95, while those who completed the high school before going to work had a median intelligence quotient of 111. Observe also that the intelligence of those who left school to go to work is 20 points below that of the group in educational institutions.

There is likewise considerable difference between those who are still pursuing their education and those who are now employed in respect to the vocational ambitions which they had four years ago. Table VII gives the details. Although scarcely anyone in either group expressed a preference for occupations of Rank IV or V, it will be seen that 44 percent of those in occupations were content with Rank III, while only 11 percent of those in educational institutions looked forward to similar vocations. The preponderance of choice for Rank I among the pupils who continued their education is also worthy of note (37 percent as compared with 20 percent).

TABLE VII. VOCATIONAL AMBITIONS OF THOSE IN OCCUPATIONS AND IN EDUCATIONAL INSTITUTIONS

	Rank IV		Rank III		Rank II		Rank I	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Group in educational institutions.....	0	0.0	32.	11.0	152.	52.0	106.	37.0
Group in occupations	1	0.4	120.	44.0	97.	35.6	54.	20.0

Summary and Conclusions

1. This study represents a fourth-year follow-up on 771 high-school pupils who had been tested for general intelligence and who had answered questions regarding vocational and educational plans in 1917-18.

2. Of these, 272 were found in occupations and 290 in educational institutions beyond the high school.

3. Among those in occupations, 40 percent were employed at tasks ranking equal to or better than their ambitions while 60 percent were in occupations of lower rank. More girls than boys were following vocations of their original choice.

4. In the same group it was found that whereas 99.6 percent had ambitions for occupations of Rank III or above, 39 percent were found in occupations of Ranks IV and V.

5. The influence of intelligence or persistence in vocational plans was found to be noticeable but not pronounced.

6. Only 26 percent of the boys and girls in occupations had had any specific training either in the high school or after leaving it for the occupations in which they were engaged.

7. There appears to be a closer relation between intelligence and persistence in educational plans than between intelligence and persistence in vocational plans.

8. In terms of median intelligence quotients, those in educational institutions surpassed those in occupations by 10 points and were superior to those who left school to go to work by 20 points.

9. The vocational ambitions of those in educational institutions ranked higher than the vocational ambitions of those now in occupations. Of those in educational institutions 89 percent had ambitions ranking I and II while only 55 percent of those in occupations had ambitions of those ranks.

10. The data presented in this study tend to support the thesis that information regarding educational and vocational plans of high-school pupils while they are still in the high school has more significance than has heretofore been attached to it, and particularly that measurements of general intelligence may be of great value to the vocational and educational counselor in his work.

11. There is also an implied indictment of the high schools studied, in the fact that although 41 percent of the cases followed up were found to have gone directly into occupations from the high school, only 71 out of the 272 had had any training in the high school for the tasks undertaken.

12. Better and more systematic educational and vocational guidance of high-school pupils, and more careful attention to those who will not go on to college, are needs that seem to be definitely indicated by the facts presented.

FINANCING SCHOOLS UNDER THE DISTRICT SYSTEM[1]

FLETCHER HARPER SWIFT
University of Minnesota

When we speak of financing schools under the district system, we are to all intents and purposes describing the system by which the public schools of the United States are supported. In the year 1920, according to data collected by the Bureau of Education, out of every $100.00 provided for public schools in the United States, $17.00 came from state sources of revenue, $11.00 from county, and $72.00 from local units. In some states the local educational unit is the town; in Louisiana it is the parish; but in the great majority of our states it is the school district. The best manner of discovering the results of the district system of school support is by turning our attention to a few individual states. In the space and time allowed, all that is permitted is a hurried view of certain outstanding facts in connection with three of several states which have been subjected by the writer to an intensive study, namely, Oklahoma, Arkansas, and Minnesota.

FINANCING SCHOOLS UNDER THE DISTRICT SYSTEM IN OKLAHOMA

In the year 1920-21, out of every $100.00 provided for public schools in Oklahoma, $1.00 came from the federal government, $9.00 from the state, $13.00 from the counties, and $77.00 from school districts. It should require no argument to secure acceptance for the statement that a satisfactory, equitable, scientific, and democratic system of school support would be one which distributes the burdens of school support according to ability. A study made of nine representative counties in Oklahoma showed that in the richest of these, where the wealth

[1] A paper read before the Administrative Research Section, Educational Research Association, March, 1923.

per child was over $10,000.00, the average rate of district tax in 1921 was 8 mills; whereas, in the poorest of the nine counties, where the wealth per child was $1,590.00, the average district tax rate was 18.8 mills. A comparison of these nine counties, county by county, showed that in general the principle holds that the richer the county, the lighter is the burden which the district system of support places upon it. The complete story is told in the following table, which shows that as wealth decreases school burdens increase.

TABLE I. HOW OKLAHOMA'S SCHOOL BURDENS WERE DIVIDED IN 1921

Group	Number of counties in group	Typical valuation per child in average daily attendance	Average district levy in mills
1	1	$7,200.00	3.7
2	6	6,500.00	5.1
3	13	5,200.00	6.4
4	7	4,600.00	9.5
5	2	4,500.00	7.4
6	8	3,800.00	8.6
7	21	3,200.00	11.8
8	7	2,600.00	10.5
9	2	2,200.00	14.5

When we turn our attention to rural districts, we find the story to be the same. The richest district was compared with the poorest in five Oklahoma counties, namely, Grant, Cimarron, Kiowa, La Flor, and Haskell. The results of this comparison are set forth in Table II.

It is impossible here to enter upon a description of the gross inequalities in educational opportunities which, in spite of the undue efforts of the poorer districts, permeate the school system of Oklahoma. These will be found set forth at length in the report, "Public Education in Oklahoma," just issued by the Bureau of Education. We may note, however, one or two pertinent facts. Twenty-six percent of Oklahoma's money goes to sixteen percent of her children. These children live in six counties. The richest district in Cimarron County, which levied a tax of only 4 mills, spent $395.00 per child in 1922; whereas,

TABLE II. WEALTH PER CHILD AND SCHOOL LEVY IN RICHEST AND POOREST RURAL SCHOOL DISTRICTS IN FIVE OKLAHOMA COUNTIES

County	District	Valuation per child in average daily attendance	School mill tax
Grant	Richest	$82,000.00	5.0
	Poorest	3,000.00	9.4
Cimarron	Richest	61,000.00	4.0
	Poorest	3,000.00	9.0
Kiowa	Richest	20,000.00	4.75
	Poorest	2,000.00	15.0
La Flor	Richest	16,000.00	12.0
	Poorest	1,000.00	15.0
Haskell	Richest	7,000.00	21.0
	Poorest	1,000.00	16.0

the poorest district, which levied a tax of 9 mills, was able to spend only $51.00 per child. The richest district in Kiowa County, which levied a tax of only 4.75 mills, spent $108.00 per child; whereas, the poorest district, which levied a tax of 15 mills, was able to spend only $28.00 per child.

Financing Schools Under the District System in Minnesota

In the year 1921, $87.00 out of every $100.00 provided for public schools in Minnesota came from the school districts, $12.90 came from the state, and $.10 from the federal government. The injustice of the district system in Minnesota is readily evident from the following facts. In the year 1919-20, districts in Minnesota paid school taxes ranging from less than 1 mill to 117 mills. In the year 1921-22, fourteen districts in the state levied taxes of more than 75 mills. Fifty-two districts levied no maintenance tax whatsoever. Most of them provided neither schools nor transportation. A 15-mill tax on their total valuation of $4,778,000.00 would have produced $71,670.00, or $1,378.00 per district.

The district system in Minnesota results in the same inequitable distribution of school burdens as that discovered in

Oklahoma. Thus in four counties in the state where the typical valuation per child was $9,250.00 the average district tax rate was 7.68 mills. In twenty-one counties, constituting a group of which the typical valuation per child was $3,250.00, the average rate of district tax was 23.4 mills. In a group composed of six counties the typical valuation of which was $2,250.00, the average rate of district tax was 26 mills. Minnesota contains eighty-six counties. Ranking the counties of Minnesota with respect to their wealth per child in average daily attendance, we find that Murray County ranks second; Sibley, sixth, Nicollet, eighth; Mille Lacs, eighty-fifth; and Hubbard, eighty-sixth. A study of the district tax rates levied within these counties shows that in Sibley County 78 percent of the districts, and in Nicollet County 52 percent of the districts levied a tax of less than 5 mills; whereas, in Mille Lacs County 95 percent of the districts levied taxes of more than 5 mills, and 33 percent of the districts levied taxes of more than 20 mills. In Hubbard County 97 percent of the districts levied taxes of more than 5 mills and 47 percent of the districts levied taxes of more than 20 mills. It is impossible to recount here the results of the district system in Minnesota. These will be found detailed in a recent brief report entitled "Public School Finance in Minnesota," copies of which may be secured by addressing Secretary C. G. Schultz, Minnesota State Education Association, 808 Pioneer Building, St. Paul. Some idea of the effects of the district system of school support in Minnesota can be gleaned from the fact that 24 percent of Minnesota's school moneys goes to 9 percent of her children. These children all live in one county, St. Louis. Comparing St. Louis County with the state as a whole, we discover the following facts. The annual expenditure per child enrolled for the state as a whole is $71.00; for St. Louis County, $181.00. The average number of days attended by each pupil in rural schools for the state as a whole is 96; in St. Louis County, 136. The average annual salary of teachers for the state as a whole is $882.00; for St. Louis County, $1415.00. The percent of teachers who are graduates of normal schools or colleges in the state as a whole is 46; in St. Louis County, 71.

Financing Schools Under the District System in Arkansas

In the year 1921, Arkansas expended for public schools $11,096,259.00. Eighty-two percent of the moneys provided for the public schools were furnished by school districts. The results of the district system of school support in Arkansas can readily be grasped by comparing certain typical counties with respect to the valuation of all school property per child enrolled, and with respect to the average annual expenditure per child enrolled. Such a comparison reveals the following facts. In the year 1921 expenditures per child varied in Arkansas from $30.54 in Pulaski County to $6.24 in Stone County. Valuation of school property per child enrolled varied from $151.00 to $11.00 in Stone County.

TABLE III. NINE ARKANSAS COUNTIES COMPARED AS TO AVERAGE EXPENDITURE AND VALUATION OF ALL SCHOOL PROPERTY PER CHILD ENROLLED

Average Expenditure per Child		Valuation of all School Property per Child	
County	Amount	County	Amount
Pulaski	$30.54	Pulaski	$151.00
Garland	18.96	Garland	72.00
Arkansas	16.34	Arkansas	56.00
White	11.19	Boone	40.00
Boone	11.12	Hempstead	30.00
Hempstead	10.21	White	28.00
Van Buren	9.42	Van Buren	21.00
Newton	8.00	Newton	19.00
Stone	6.24	Stone	11.00

A study of these same nine counties with respect to the length of the school year shows that in Pulaski County the average length of the school year was 168 days; in Arkansas County, 147 days; Garland, 142 days; White, 122 days; Hempstead, 117 days; Boone, 109 days; Stone, 89 days; Van Buren and Newton, 77 days. The extremes of inequalities produced by the district system are perhaps nowhere better shown than in a study of school districts in Arkansas. For purposes of comparison it was de-

cided to select the average monthly salary paid to women in rural common school districts; the length of the school year in months, weeks, and days; and the average annual expenditure per child. In a state where the negro problem looms so large, it is necessary to separate white from colored schools when such comparisons are made. The above comparison showed that the average monthly salary to women teachers in white rural schools varied from $114.00 in Little River County to $53.00 in Izard County, and in negro rural schools from $75.00 in White County to $35.00 in Perry County. The school year in months, weeks, and days varied in white schools from 8 months and 2 weeks in Pulaski County to 3 months, 3 weeks, and 2 days in Newton and Van Buren counties; in negro schools from 8 months and 2 days in Hot Springs County to 3 months in Izard County. The average annual expenditure per white child enrolled varied from $46.22 to $3.75 in Maria County. The highest average annual expenditure per negro child was $14.45 in Garland, and the lowest was $2.20 in Johnson County.

It is impossible in the present brief presentation to enter upon any full consideration of the causes of the inequalities produced by the district system. Staunch defenders of the district system commonly attempt to support their claims for it on the ground that it is democratic, that it encourages local initiative, and that it educates the people. The essence of democracy is equality—equality in responsibility and in opportunity. School districts wherever found are so unequal in wealth—that is, in ability to provide school revenue, in intelligence, and in zeal for education that the district system never has and never will produce equality in any aspect of education. A system more pernicious, more unworthy of an intelligent commonwealth could scarcely be devised. It has been condemned for generations by all important leaders in educational administration and by students of school finance. It is the cause, the mother, the nurse, and entrenched dowager of injurious vested rights. Wherever found it produces the same evils—financial and educational inequalities and injustices as incurable as they are numerous and disastrous.

Brief as is the space allotted to this presentation, it would seem unpardonable not to call attention to the fact that the district system as a rule depends for the greater part of its revenue upon the levy of a general property tax. States which are overwhelmingly rural must undoubtedly continue to derive the major proportion of their public revenues from a general property tax. Nevertheless, as rapidly as possible other sources should be provided. The necessity of this has been pointed out both by individual students of taxation and by public commissions. Professor Seligman of Columbia University in "Essays In Taxation," page 62, writes, "The general property tax is one of the worst taxes known in the civilized world." The Georgia special tax commission in its report to the Governor in 1919 reported that the general property tax is a failure, and recommended taxes on incomes, inheritances, privileges, and occupations. The Ohio Tax Commission declared that the general property tax punishes the honest, rewards the dishonest, often results in double taxation, and lowers the standard of integrity. Elsewhere I have discussed the remedy for the district systems,[2] and the part that the state must play in the support of public schools.

Evidences are not wanting that many of our states are rapidly coming to realize the evils of the district system and the necessity of equalizing school burdens and educational opportunities by deriving revenue from larger units, notably the counties and the state. California in 1921 increased her state aid per child from $17.50 to $30.00 for elementary schools and from $15.00 to $30.00 for high schools. She now guarantees from state and county sources together an annual minimum of $1400.00 for every full-time elementary teaching position. Minnesota at the present time has before her legislature a bill providing for a permissive county school tax. Many other examples might be given. It is to be hoped that such sound tendencies will rapidly gain headway; but let those who are interested in bringing about reforms in school finance be fully conscious of the organized

[2] Swift, F. H. *State policies in public school finance*, Bureau of Education Bulletin, 1922, No. 6. 54 pp.

effort which will be necessary to stem the dominant and dangerous trend represented by the fact that, whereas in 1890 the public schools of the United States derived 76 percent of their total revenue from local sources[3] (chiefly the school district), in 1905 they derived 80 percent and in 1920, 83 percent from this source. Throughout the history of the last thirty years the almost universal tendency has been for the state to shirk more and more its share of the burden of school support and to place an ever-increasing load upon the weakest and the most unequal units at its command, namely, the school districts. This situation must be remedied and our national policy reversed if the people of the United States are to equalize school revenues and school burdens and provide equal educational opportunities for the children of the United States.

[3] Includes receipts from miscellaneous sources.

THE ANALYSIS OF TEACHING FUNCTIONS

FREDERICK L. WHITNEY
State Teachers College, Moorhead, Minnesota

Administrators of teacher-training institutions need tools and methods of using them so that, with the material they have at hand, they may proceed more surely in the prosecution of their task of making public school teachers from the raw material coming from the secondary schools. It is the purpose of this article to present two devices which may be used with the data usually available in any state normal school or teachers' college. These devices depend upon the method of partial correlation. They are the coefficient of multiple correlation (R) and the regression equation.

The present study deals with nine measurements of traits obtained from 1200 graduates (classes of 1920) of two-year college courses in twelve representative state normal schools. These are (1) teaching success, as judged by supervisory officers during the first semester under contract after graduation; (2) intelligence, as measured by tests for general intelligence such as the Army Alpha and Thorndike's Intelligence Examination for High-School Graduates; (3) secondary record, an average of all high-school marks earned during a four-year course; (4) academic record, an average of all marks in normal-school subject-matter courses; (5) professional record, an average of all marks in normal-school courses in education, educational psychology, and the like; (6) student teaching success, an average of ratings in practice teaching made by the training school faculty; (7) physique, a checking of such items as chest expansion, posture, lung capacity, weight, general health, etc.; (8) chronological age, number of years since birth; (9) personality rating as given by the normal-school faculty. (Original measures are on file in the office of the Graduate School, University of Minnesota.)

For purposes of illustrating the method of analysis here given, the first seven of these available items have been chosen. Chrono-

logical age and personality ratings are omitted because in the case of chronological age the coefficients of correlation of order zero are negative or very low, as is usually found in such relationships, and because in the case of personality, data from but two of the cooperating schools are available, and these are of doubtful value. It was also found to be necessary to reduce the number of variables used so that, in the time available, all necessary computations could be completed. About 275 coefficients of order zero and about 800 of higher orders were computed before either R or the elements of the regression equation could be calculated.

The method of treatment involved first a discovery of relationships among the seven variables used. The median coefficients of order zero were found to be as shown in Table I. The digits refer to the seven traits (variables) as numbered above.

TABLE I. COEFFICIENTS OF CORRELATION (ZERO ORDER) BETWEEN EACH PAIR OF TRAITS

	1	2	3	4	5	6
2	.025					
3	.091	.314				
4	.073	.343	.355			
5	.143	.327	.355	.484		
6	.238	.160	.266	.386	.207	
7	.124	.201	.192	.099	.078	.303

Table I reads: r_{12}, or the coefficient of correlation between (1) teaching success and (2) intelligence, is 0.025; r_{13}, or the coefficient of correlation between (1) teaching success and (3) secondary record, is 0.091, etc. These coefficients, however, are thought to be but crude expressions of the relationships to be studied. It is difficult to interpret, for example, a coefficient of correlation of order zero. What does $r_{16}=.238$ mean? We cannot know except through a vague guess. And this will be a guess on the basis of very meager data. If teaching success after graduation and student teaching performance correlate to the degree 0.238, it may be said that scores in the complex activity

"real teaching" correlate positively with markings in the complex activity "student teaching," but this really does not mean very much since we have no statement as to what these two complex functions include. It is similar to saying that the inhabitants of Mars, if there are any, are very like the people of the moon, if it has any.

But, to change the variables used, suppose one should consider the relation of (1) teaching success to (2) intelligence and to (3) secondary record. Calculation shows the first (Table I) to be r_{12}=.025, and the second to be r_{13}=.091. Here more facts appear for use, and one is tempted to try to argue about what is measured by r_{12} and r_{13} without further data. But common sense will not let us assume that any two variables can operate independently in a complex realm of relationship. There are always many other factors whose influence must be accounted for. The method of partial correlation makes possible the consideration of as many of these factors as we are able to use in terms of expressions of amount.

Partial correlation, then, dealing with coefficients of higher orders, helps to put content and meaning into expressions of relationship between two related activities. For example, to return to the consideration of the factors (1) teaching success and (6) student teaching, calculation shows that $r_{16.2}$=.236.[1] This may be read, "The correlation between teaching success and student teaching *with intelligence kept constant* is positive

[1] The reader may not be familiar with the usual notation for designating coefficients of partial correlation. Relationship is always expressed between two and only two variables. One or more other variables, however, may be "counted out" or made constant. Assuming, as is the case in this article, that the variables are numbered, the symbol r is written with the numbers of the variables as subscripts. The two between which the relationship is expressed are written first; then a point is inserted; and after the point are written the numbers of the variables which are counted out. For example, r_{12} expresses the gross relationship between the first and second variables. This is a correlation of order zero. Again, $r_{12.3}$ expresses the same relationship with the third variable counted out or constant. This is a correlation of the first order. Similarly, $r_{12.34}$ counts out the third and fourth variables and expresses a correlation of the second order. Thus the order of the correlation is shown by the number of digits to the right of the point. Other constants such as the regression coefficient (generally designated by b) and the root mean square error (s or σ) are expressed with similar subscripts having like meanings.

and to the degree expressed by the ratio 0.236." That is, an answer appears here to the question, "What in student teaching is related to teaching success that is peculiar to it and has nothing to do with general intelligence as measured by the particular tests used?" If $r_{16.2}$ were zero, then it could be said that everything in intelligence which might be associated with teaching success is found in student teaching. But in proportion as $r_{16.2}$ approaches 1.00 (perfect correlation), the influence of the variable "intelligence" affects the relationship of teaching success and student teaching increasingly.

Analysis of relationships may begin on the level of zero calculations, when but two variables are accounted for; but it will always be with insufficient data and can lead to but very vague conclusions. For example, about all that can be said concerning the relationships r_{12} or r_{13} given above is that in either case the activities are both alike and different, and that this would depend upon what aspects of the complex functions one had in mind.

But the consideration of a third factor in the realm of concomitance makes analysis on a higher level entirely possible. And further calculation with a fourth variable points to a second step. It is found, for example, that $r_{16.23}$=.222. This may be read, "The relation of student teaching to teaching success, *intelligence and secondary record kept constant,* is positive and to the degree expressed by the decimal 0.222." Moreover, it is now evident that "secondary record" was contributing something to the relationship; for before it was isolated and kept constant, $r_{16.2}$ was 0.236,—a higher value than 0.222. Further, it is found that $r_{16.234}$=.221; that $r_{16.2345}$=.192; and that $r_{16.23457}$=.195. Since the addition of each successive variable to those which are kept constant changes r_{16} but little, we may venture to say that there is nothing in any of these variables that affects very much the relation between teaching success and student teaching. Perhaps it may be added that professional marks and physique affect it not at all.

The thirty-six combinations of variables, showing the relation of the criterion (i.e., the dependent variable, teaching suc-

cess) to each independent variable when the influence of the other variables in increasing groups is kept constant, are as follows. In every case, calculations are with raw zero r's. That is, the r's of zero order which are used in deriving the r's of higher order were not corrected for attenuation.

Group I — r_{12}

(a) r_{12}	.025
(b) $r_{12.3}$	—.033
(c) $r_{12.34}$	—.014
(d) $r_{12.345}$	—.033
(e) $r_{12.3456}$	.029
(f) $r_{12.34567}$	.022

Group II — r_{13}

(a) r_{13}	.091
(b) $r_{13.2}$	.086
(c) $r_{13.24}$	.069
(d) $r_{13.245}$	.022
(e) $r_{13.2456}$	—.030
(f) $r_{13.24567}$	—.036

Group III — r_{14}

(a) r_{14}	.073
(b) $r_{14.2}$	.067
(c) $r_{14.23}$	.044
(d) $r_{14.235}$	—.063
(e) $r_{14.2356}$	—.118
(f) $r_{14.23567}$	—.112

Group IV — r_{15}

(a) r_{15}	.143
(b) $r_{15.2}$	.141
(c) $r_{15.23}$	.121
(d) $r_{15.234}$	.114
(e) $r_{15.2346}$	.099
(f) $r_{15.23467}$	.101

Group V — r_{16}

(a) r_{16}	.238
(b) $r_{16.2}$	.236
(c) $r_{16.23}$	.222
(d) $r_{16.234}$	.221
(e) $r_{16.2345}$	.192
(f) $r_{16.23457}$	.195

Group VI — r_{17}

(a) r_{17}	.124
(b) $r_{17.2}$	.121
(c) $r_{17.23}$	.110
(d) $r_{17.234}$	.014
(e) $r_{17.2345}$	.020
(f) $r_{17.23456}$	—.037

In general, the coefficients seem to show that there is something in each of the independent variables—i.e., variables (2) to (6)—which affects to some degree the relationship of the criterion with each remaining variable. This view is supported by the fact that in general the coefficients in every group decrease to some degree as the influence of each remaining variable is successively removed. Perhaps the variable "professional record" (number 5) is the most influential of all since in four cases out of five (Groups I, II, III, and V) the coefficient is

very materially reduced in value when this is used in the calculation of the partial coefficient of correlation.

When one compares items "a" and "f" in each group above, one observes that the "a" items are larger. It is quite apparent that coefficients of order zero are not true expressions of real relationship. Something is evidently contributed to them from at least five other items of environment in each case. In fact, the "f" coefficients represent in each case the correlation with teaching success, the criterion, of what is unique to each of the independent variables.

Again, if the coefficient r_{12}=.025 is considered, it cannot be said that this is really a true measure of the degree to which intelligence is a factor in teaching success, for (a) intelligence as measured by the Army Alpha may not be a very good measure of teaching success because of lack of similar elements in the latter activity, and (b) it may be that these intellectual elements are not really the vital elements either in general intelligence or in teaching intelligence. The low raw coefficient above[2] seems to bear out this viewpoint. The complex teaching function certainly is not confined to activities represented by the eight elements of alpha.

Now, if it were possible to find a better index for teaching success, one might proceed to determine true relationships more successfully. If (3) secondary record, for example, might be thought of as more nearly like teaching success, then $r_{12.3}$ would carry a stronger hint of the real teaching function, and $r_{13.2}$ might point more toward intelligence. In each case, the common element has been dropped, and what is inherent in teaching success becomes more apparent. The following group of coefficients illustrates all possibilities of this type.

(a) $r_{12.3}$ = .120	(a) $r_{13.2}$ = .254
(b) $r_{12.4}$ = .124	(b) $r_{14.2}$ = .223
(c) $r_{12.5}$ = .068	(c) $r_{15.2}$ = .321
(d) $r_{12.6}$ = .098	(d) $r_{16.2}$ = .461
(e) $r_{12.7}$ = .146	(e) $r_{17.2}$ = .317

[2] In the larger study of which this is a partial report, a criterion is established for attenuation of zero coefficients amounting to about 0.30. Calculations on this basis are carried parallel to figures based on raw coefficients, but these are not included in this brief report.

These combinations of variables in relation to a criterion might be carried conceivably to curious limits involving all the possible combinations of seven things. One group of forty-two coefficients is reported in the original material of this study, but is not reproduced here. Many significant facts appear in this array of relationships. For example, it is noticed that on the whole, no matter which variable is taken as the criterion, the coefficients decrease in size as more independent variables are added to the "constant" group. And it may be assumed that the same interpretation can be given for this fact as was given for a similar fact above where teaching success was used invariably as the criterion.

It would be interesting, too, to follow each of the variables through the seven groups of six coefficients each and to discover its value in every group in which it is found. For example, teaching success, when used as the criterion with each independent variable successively and all remaining variables constant, has an average r of —0.117 (raw). Evidently, there is something in each independent variable pertinent to teaching success which should be considered in preparation for performance in actual teaching under contract.

The intelligence group is as follows:

(a) $r_{27} = .201$
(b) $r_{26.7} = .105$
(c) $r_{25.67} = .308$
(d) $r_{24.567} = .216$
(e) $r_{23.4567} = .159$
(f) $r_{21.34567} = .022$

Here, the criterion exhibits the same type of relationship to remaining variables. And it would seem that we must conclude that it is but the sum of partial intelligences all of which enter into a composite "teaching intelligence."

But space cannot be used to discuss the many implications which appear. All that can be done is to show some of the possibilities in the method of partial correlation as a means of analyzing complex functions such as those measured in this study.

The Coefficient of Multiple, or *n*-fold, Correlation

Another method of analyzing variables is by means of the coefficient of multiple correlation. Yule[3] has suggested the notation "R" for this coefficient, and has deduced the formula for its calculation,[4]

$$R_{1.234567} = [1-(1-r^2_{12})\ (1-r^2_{13.2})\ (1-r^2_{14.23})\ (1-r^2_{15.234})\ (1-r^2_{16.2345})\ (1-r^2_{17.23456})]^{\frac{1}{2}}$$

In the same manner that the coefficient of correlation of order zero is an expression of the value for prognosis of one variable in terms of another, so R is the correlation of a group of independent variables with a criterion. But for the purposes of this section, its value as a tool for analysis of functions is the topic. R for both raw and corrected data has been calculated, but the latter cannot be given here nor the details of the computation for the former. The coefficients of multiple correlation are given below not only for the criterion and all six variables, but when each of the variables is successively dropped from the independent group.

Groups	Coefficients
All items	
$R_{1.234567}$	.288
Without intelligence	
$R_{1.34567}$	.266
Without secondary record	
$R_{1.24567}$	.264
Without academic record	
$R_{1.23567}$	.242
Without professional record	
$R_{1.23467}$	.246
Without student teaching	
$R_{1.23457}$	.152
Without physique	
$R_{1.23456}$	.242

[3] Yule, G. U., *An introduction to the theory of statistics*, London, 1919, p. 248 ff.

[4] Subscripts when attached to R have the same appearance as those attached to r; but they do not have the same meaning. R is a measure of the correlation of one variable with a *group* of others taken as one. In the writing of the subscripts the one variable (or the criterion) is given first, then a point is inserted, and after the point is set down the group of variables which when taken together constitute the multiple unit with which the criterion is related.

Here analysis on the basis of R begins to be possible, for it is noticed that the relationship of success in teaching to the group of variables is different when different members of the group are omitted. The most marked reduction from 0.288, the total coefficient, occurs when student teaching is not found among the group of variables; and we might conclude from these figures that, as preparation for teaching success after graduation, student teaching is the most important single item of experience.

The variables seem to rank as follows, in their relationship to teaching success: (a) student teaching, (b) academic records, (c) physique, (d) professional record, (e) secondary record, and (f) intelligence.

In a battery of independent variables, then, one measure of their relative importance may be found in R, the coefficient of multiple correlation. This is made clear if R is computed first for the criterion as related to the entire group and then for the criterion as related to each group when the independent variables are discarded successively.

The next group of coefficients results from every possible pairing of the six independent variables, R being computed for each case.

$R_{1.23}$=.028 $R_{1.24}$=.100 $R_{1.45}$=.143 $R_{1.56}$=.252 $R_{1.67}$=.243
$R_{1.24}$=.022 $R_{1.35}$=.148 $R_{1.46}$=.239 $R_{1.57}$=.182
$R_{1.25}$=.143 $R_{1.36}$=.239 $R_{1.47}$=.138
$R_{1.26}$=.237 $R_{1.37}$=.141
$R_{1.27}$=.123

Analysis is possible here also, for to a constant independent variable (e.g., in column 1) is added successively each remaining independent variable, and the relationship is shown to differ with each addition.

Considering (6) *student teaching* first of all, in each group its addition increases the value of R more than does the addition of any other variable and the coefficients in which it appears are the largest in the table.

Considering the first group of five R's above, the order of importance of each independent variable outside of intelligence

is: (a) student teaching, (b) professional record, (c) physique, (d) secondary record, and (e) academic record. In the second group of four, the order is: (a) student teaching, (b) professional record, (c) physique, and (d) academic record. Similar statements may be made from the data of the third and fourth groups.

When all rankings are summated, the following order is found: (a) student teaching, (b) professional record, (c) physique, (d) academic record, (e) intelligence, and (f) secondary record. In fact, the analysis by means of R is remarkably consistent, and it may be that we have here one good method of evaluating the different items of personal equipment and preparation which are employed in attempting to insure the teaching success of prospective teachers. Other groupings of the independent variables cannot be reported here. Enough have been given to illustrate the method proposed, and further analysis but confirms the above findings.

The Regression Equation

The coefficient of multiple correlation, R, has been shown to act as a tool for the analysis of a complex situation not alone through the discovery of comparative values among the different members of the correlating group but by means of the evaluation of an entire independent group of correlating factors in terms of a criterion. In like manner, the regression equation may be used to discover proper weights to be assigned to the separate factors of a correlating group, weights which will correspond to their relative influence in the realm of concomitant variation under consideration.

Knowledge of the detailed process necessary in writing the regression equation must be assumed here. When but two variables are used its theory is as simple as the well-known equation, $v = gt$. The consideration of more factors in the realm of correlation studied requires only that more elements be expressed and summated at the right side of the equation. For

seven variables, it becomes, in terms of deviations from a central tendency,[5]

$$x_1 = b_{12.34567}x_2 + b_{13.24567}x_3 + b_{14.23567}x_4 + b_{15.23467}x_5 + b_{16.23457}x_6 + b_{17.23456}x_7$$

For the particular realm of influence here studied, it will be expressed as $x_1 = .004x_2 - .038x_3 - .115x_4 + .915x_5 + 1.403x_6 - .085x_7$. In other words, the following weights must be assigned to the six independent variables in order to predict accurately performance in teaching success; (2) intelligence, 0.004; (3) secondary record, 0.038; (4) academic record, 0.115; (5) professional record, 0.915; (6) student teaching, 1.403; and (7) physique, 0.085.

When these factors are arranged in rank order, they become: (a) student teaching, (b) professional record, (c) academic record, (d) physique, (e) secondary record, and (f) intelligence.

This confirms, on the whole, the analysis by means of R above. Physique and academic record have changed places as well as intelligence and secondary record, but student teaching and professional courses remain at the top of the column as preponderantly of greatest weight in the preparation of elementary teachers.

Analysis of this sort might be continued, of course, by a rearrangement of all seven variables so that each in turn would become the criterion in the writing of the regression equation. One such calculation will be reported in which the criterion is taken to be (6) student teaching, and the independent variables, (2) intelligence, (3) secondary record, (4) academic record, (5) professional record, and (7) physique. Here, the proper weights for the five factors affecting student teaching success are found to be: intelligence, 0.053; secondary record, 0.014; academic record, 0.382; professional record, 0.098; and physique, 0.093.

[5] Where x_1 is the deviation of teaching success, x_2 that of the intelligence score, etc.; and where the b's are the coefficients of regression. The subscripts written in connection with the b's have reference to the variables in the same way that they do in partial correlation when written in connection with the x's.

It will be observed that the order of importance of these variables is practically the order found for them in respect to their influence upon teaching success after graduation. This confirms the value of the regression equation as a tool for analyzing correlation in a given realm of complex concomitant variation. In fact, this calculation reveals nothing entirely new with regard to relative weights, and very probably further computations of like nature would not either. They would, however, be a check upon the calculation for R, the coefficient of multiple correlation, and for this reason might be worth while.

Summary

The three methods of analysis illustrated above are all possible only through the use of the method of partial correlation. First, the partial coefficients are used directly. By means of comparisons of values and by addition and subtraction of different members of the correlating group, many valuable facts are revealed. Next, R, the coefficient of multiple correlation, which is obtained through the agency of partial correlation coefficients, analyzes the combined value of the group of independent variables and, when different parts of the group are introduced, discovers added facts with regard to the value of individual members. Last, the method of partial correlation is used again, and a definite statement of weights in terms of amount is discovered through the writing of the regression equation.

Since this is a statement of the value of tools and methods only, the many implications for teacher-training administration suggested by the illustrations given cannot be discussed, but must be left for another article.

A SURVEY OF THE TEST MOVEMENT IN HISTORY

PAUL TYLER KEPNER
Head of History Department, High School, Taunton, Massachusetts

Nothing has been accomplished as yet in the preparation of tests for the field of social sciences as a whole. So far as one is able to learn, progress has been made in only one part of the field, namely, history. One is able to find 22 tests or scales for history in secondary schools, prepared by 16 different agencies. Of these 22 tests, 2 are for use in the field of ancient history, 15 for American history, and 5 for the testing of general ability in history. At the time of writing, 13 are known to be available for use in prepared forms.

The most general criticism that can be passed upon the group as a whole is the fact that 15 tests, or more than half, aim to test only one ability, that is, are only informational in character. These tests have, therefore, not kept pace with the newer methods of teaching in the field of history. Teachers in increasing numbers are coming to recognize that historical information per se is but a means and not an end. It is only too well known, as results have shown, that pupils do not retain much historical information. Bell and McCollum have shown experimentally that "simple questions on fundamental aspects" of United States history are answered 33 percent correctly by high-school pupils.[1] There are other abilities to be tested in history, which, in the minds of progressives, are of greater value.

History, in so far as the measurement of results is concerned, is unlike the more objective studies. It has been difficult to determine a test satisfactory to any great number. Even in the field of information alone agreement is difficult, for the field is unlimited in extent. Content vital to one course of study may not be thought so for another. The recent shift of emphasis, evident in texts, has interrupted any progress that may have been

[1] "Study of attainments of pupils in United States history," *Journal of Educational Psychology*, 8:257-74, May, 1917.

made in the standardization of content. The majority of tests, as has been shown by E. U. Rugg in the best treatment that has as yet appeared, do not test "basic aims and outcomes" of history.[2] But the primary difficulty in devising history tests is the fact that the valid aims and outcomes of history have not as yet been experimentally determined.[3] If, however, recent changes and tendencies of recognized texts in the field are to be permanent, greater emphasis on content in the future will be placed upon recent periods of history, and if recent tendencies in the presentation of the subject are generally accepted, emphasis in aims in the future will be placed not upon historical information but upon ability to use historical material—that is to set at work the mental processes of thought, reasoning, inference, and judgment, and imagination to some extent. The greater number of history tests appeared before recent changes and tendencies in the field crystallized.

The tests now in the field will be approached critically where available material permits, from the point of view of the essentials of a test, such as source, basis of standardization, type, kind of questions, available series, scoring, and time of administration. If details seems rather numerous in some instances, it is because the writer believes that the intrinsic value of any kind of test, within or without the field of history, rests largely upon the precise methodology employed in arriving at the finished product. For purposes of convenience and summary Table I, based upon material of available sources, has been prepared.

Sources of Tests

In originating a test, the author has in general three paths which he may follow. He may choose to base his test upon the dicta of "competent judges" or he may, as some have done, place it upon the foundation of his own personal judgment or

[2] "Character and value of standardized tests in history," *School Review*, 27:757-71, December, 1919.

[3] For the psychologist's defense, or attack as the case might be, *cf.* H. O. Rugg, "How shall we reconstruct the social studies curriculum?", an open letter to Professor Henry Johnson, *Historical Outlook*, 12:184-89, May, 1921. A second article appears in the October, 1921 issue, 12:247-52.

he may proceed experimentally. He may, also, use all approaches and give his test a triangular basis. Of this latter type are the tests of Bell and McCollum, and Buckingham. As indicative of the empiricism employed in the creation of tests, several assumptions, which in part underlie Buckingham's questions, are given: (1) "the more difficult a question is the more ability does a child show in answering it," (2) "that a scale derived from memory questions might serve to measure the thing we are calling ability in history."[4] On one question 100 competent judges passed upon the five most important events included. Sackett's ancient history scale is another example of the threefold approach, but bears the serious limitation of an exceedingly narrow selection of material inasmuch as the content is taken wholly from one text, Meyers' "General History."

The tests of Barr, Davis, Harlan, and Rayner[5] are based in whole or in part for content on the study of 23 textbooks made by Bagley and Rugg to determine the content of American history. In addition to the above basis, it is evident, according to Griffith, that Harlan must have "decided upon questions that range from easy to difficult, that touch upon the several different phases of history, that can be briefly stated, that afford answers easily scored as right or wrong, and that can be completed by all pupils in the ordinary history period."[6] The distribution of questions over the periods into which American history divides itself, is determined in the Rayner test by the space devoted to those periods in seven texts published since 1904.[7] Further reliance is placed upon competent judgment in the selection of dates, which is based largely upon the study of Bagley.[8] Sackett tells

[4] Buckingham, B. R., "Correlation between ability to think and ability to remember, with special reference to United States history," *School and Society*, 5:443-49, April, 1917.

[5] Bagley, W. C., and Rugg, H. O., *The content of American history as taught in the seventh and eighth grades*, Bulletin, No. 16. Urbana, Ill.: University of Illinois, School of Education.

[6] Griffith, G. L., "Harlan's American history tests," *School Review*, 28: 697-708, November, 1920.

[7] Mowry, Woodburn and Moran, McMaster, Channing, Morris, Turpin, and Meany.

[8] Bagley, W. C., "The determination of minimum essentials in elementary geography and history," *Fourteenth Yearbook of the National Society for the Study of Education*, Part I, 131-46.

TABLE I. HISTORY TESTS

Name of Author	Title	Where Obtained	Type	Time (minutes)	Basis	Series
Ancient History						
Barnard, A. F.	Roman History Test	Author, University High School, University of Chicago, Chicago, Ill.	I-C	.....	....	.....
Sackett, L. W.	Ancient History Scale	Author, University of Texas, Austin, Texas.	I	30–45	i	.....
United States History						
Barr, A. S.	Diagnostic Tests	Public School Publishing Company, Bloomington, Illinois.	I-T-R-J	35	i	2A & 2B
Boston[b]	Research Tests	Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.				
	a. Grade VI		I-C	.....	903	.....
	b. Grade VII		I-C	.....	993	.....
	c. Grade VIII		I-C	.....	917	.....
Davis, S. B.	U. S. History Exercises (Colonial Period)	Author, University of Pittsburg, School of Education, Pittsburg, Pa.	I	12	....	A & B
Hahn, H. H.[a]	History Scales	Author, Wayne Normal School, Wayne, Nebraska.				
	a. Grade VII		I-C	u	i	.....
	b. Grade VIII		I-C	u	i	.....
Harlan, C. L.[a]	Information Test	Public School Publishing Company, Bloomington, Illinois.	I-C	35	2000	.....
Rayner, W. H.[b]	American History Test	Bureau of Educational Research, University of Illinois, Urbana, Ill.	I-C	u	....	.....
Sackett, L. W.[a]	U. S. History Scale	Author, University of Texas, Austin, Texas.	I	40	1500	.....
Spokane (Wash.)	U. S. History Scale	History Department, Lewis and Clark High School, Spokane, Wash.	I-C	.....	....	.....
Starch, D.[a]	American History Test	University Cooperative Company, 504 Starr St., Madison, Wis.	I-C	u	2000	A, (B, C, D)
Van Wagenen, M. J.[a]	American History Scales	Bureau of Publications, Teachers College, Columbia University, New York City.				
	a. Information		I	45–50	2350	A & B
	b. Thought		T	45–50	2050	A & B
	c. Character Judgment		X-J	45–50	2000	A, B & L

	b. Thought	New York City.	T	45–50	2050	A & B
	c. Character Judgment		X-J	45–50	2000	A, B & L

TABLE I. (*Continued*)

Name of Author	Title	Where Obtained	Type	Time (minutes)	Basis	Series
General Ability in History						
Bell, J. C. & McCollum, D. F.[a]	History Tests	*Journal of Educational Psychology*, 8:257-74, (May, 1917).	I	30	1500	.....
Buckingham, B. R.[c]	History Tests	*School and Society*, 5:443, (April, 1917).	T	.....	544	A & B
Kelly, T. L.[b]	History Test	Teachers College Contribution to Education, No. 71, Teachers College, Columbia Univ., N. Y.	T-J-I-A	30	....	.....
Rugg, E. U.[b]	Historical Judgment Tests	Author, Lincoln School, 646 Park Ave., New York City.	J	.....	i	.....
Theisen, W. W.[d]	General History Test	The Parker Co., Madison, Wis.				

[a]Standardized.
[b]Not available.
[c]Standardized but not available.
[d]Now in preparation.

NOTES: A—Appreciation. I—Information. C—Completion. T—Thought. R—Reasoning. J—Judgment. X—Character. u—Unlimited. i—Indefinite.

us that his American history scale is based upon the tabulated data of Bell and McCollum.[9] The Starch test, Dr. Starch tells the writer, was based upon the content of five widely used texts, which furnished the basis of some three hundred statements used as working capital. The published results of Van Wagenen's investigations (investigations leading to a Columbia Ph.D. thesis, 1916) do not tell us how he determined upon the content of his questions. We do know, however, that whatever was the source of the questions, the selection in the reduction of the longer lists was approached in a thoroughly scientific manner.[10]

Standardization of Tests

Formal tests for the purpose of meeting administrative demands are, however, rapidly losing favor. In general the active and observant teacher finds no necessity for testing at the close of a period in order to determine a given pupil's grade. Tests to check the teacher's, not the pupil's, efforts are, on the other hand, very necessary. The degree of success with which teacher A is instructing a group of 150 pupils in United States history depends not so much upon what a similar group accomplished in the same school in a preceding year, but what similar groups of pupils are achieving in other schools. The standardized history test makes such comparison possible for it has established norms which represent the achievements not of a group of 150 pupils, but of many times such a group, not of a school of one locality, but of many localities and sections. Moreover, teacher A may find that while his pupils are showing excellent progress in factual history, they are showing a marked inability to use the facts of history. The diagnosis is made; the application of the remedy lies with the teacher.

It can be seen that a history test which has been carefully standardized has a wide use, and, by the same token, a test which

---

[9] Sackett, L. W., "A scale in United States history," *Journal of Educational Psychology*, 10:345-48, September, 1919.

[10] Van Wagenen, M. J., *Historical information and judgment in pupils in elementary schools*, (Teachers College Contributions to Education, No. 101.) New York: Teachers College, Columbia University, 1919. 74 pp.

lacks standardization, however excellent it may be in other respects, is distinctly limited in its utility. An examination of Barr's Diagnostic Tests indicates a series of tests carefully prepared and arranged. The most serious and far-reaching criticism of these tests, however, is that, although they have been three times revised, they are still in the process of being standardized. It is only fair to point out here that the aim of these tests is primarily diagnostic and only secondarily do they measure achievement.

The standardized tests, as Table I indicates, are those of Harlan,[11] Hahn, Sackett (United States History Scale), Starch, Van Wagenen, and Bell and McCollum. The Buckingham test must be added with considerable reservation if by standardization one means the establishment of norms based upon a fair number of cases. Information regarding the standardization of these tests is shown in Table II.

TABLE II. STANDARDIZATION OF HISTORY TESTS

Test (Identified by Author)	Basis of Standardization
Harlan	1200 VII-, VIII-grade pupils 600 high-school and normal-school students Revised by 2000 answers to each question (23 schools in 9 cities)
Hahn	43 schools in 14 states
Sackett	1500 pupils from grade V to university seniors
Starch	2000 cases (tentative norms)
Bell and McCollum	1500 papers (5 elementary and high schools and 2 normals, Texas)
Buckingham	1915—159 grade-VIII pupils (New York) 1917—288 grade-VIII pupils (Wisconsin)
Van Wagenen	1916—Preliminary tests given to 1200 pupils 1917—Information scales 2350 pupils Thought scales 2050 pupils Character scales 2000 pupils

Types of Tests

There are two distinct types of history tests, that which tests historical information per se, and that which tests historical abilities other than the ability to acquire and retain information

[11] Harlan, C. L., "Educational measurements in the field of history," *Journal of Educational Research*, 2:849-53, December, 1920.

as such. The informational tests, as Table I indicates, are those of Barnard, Sackett, Boston, Davis, Hahn, Harlan, Rayner, Spokane, Starch, Van Wagenen (Information Scale), and Bell and McCollum. The Barr and Kelly tests are partially informational.

The usual criticism made of the informational tests is that they require a knowledge of content and material no longer taught, or facts which in themselves are unimportant. The former criticism, of course, originates with that group which has allied itself with the newer tendencies in history teaching. The informational tests were constructed for the purpose of testing historical information and its retention and nothing else directly. Critiques, in and out of print, should be evaluted with this fact in mind. To the criticism that many facts tested in the informational tests are unimportant, one can only point out that this or that question which seems unimportant from the individual teacher's point of view may not be so from another point of view. Since content has never been standardized, is A to call B's judgment at fault? Such a criticism is less likely to come from one teaching informational history.

Tests aimed to test mental abilities other than the retention of information wholly are the tests and scales of Barr (thought, reasoning, and judgment), Van Wagenen (thought and character judgment), Buckingham (thought), Kelly (thought, judgment, and appreciation), and Rugg (judgment).

Buckingham and Van Wagenen have shown that there is a fair degree of correlation between ability to retain information in history and ability to use the facts of history in historical judgment, reasoning, thought, and appreciation. It will be recalled that this was one of the two basic assumptions made by Buckingham when originating his test. The correlation between ability to think and ability to remember in history was found by Buckingham to be about 0.4, "with a probable error equal to about one-tenth of its size." Van Wagenen gets amazingly high correlations. For example, he finds the correlation between information and thought to be 0.8, between thought and character

judgment to be 0.78, and between information and character judgment to be 0.7.

Tests which are in whole or part completion tests are the Barnard, Boston, Hahn, Harlan, Rayner, Spokane, and Starch scales. The following example of a completion test question is taken from Rayner:

35. In the early election in the United States the . . . . Party was unfriendly towards popular government.

The Spokane test in full and the Harlan test in part do not leave the completion entirely to the rote memory of the pupil, but furnish possible answers from which the correct answer for the uncompleted part or parts is to be selected. In parts the Van Wagenen Informational Series has a similar arrangement. While the element of chance thus enters, it can not be gainsaid that elements of selection and reasoning to a certain degree also enter. In general the kinds of questions which appear in the tests are those that may be classified as men-events, events-men, place-events, events-years, dates-events, approximate dates, terms-examples, problems-solutions, periods-men, movements-significance, causes-results, and historical significance. Examples of only the strongest and weakest features of the tests can here be given.

Several of the tests may be questioned on the grounds that their particular form necessitates the testing of abilities other than historical ability. An outstanding example of such a defect is the Kelly test.[12] This particular test is weakened in its prognosis of historical ability by reason of the many distracting elements, one of which is English. To a lesser extent the character-judgment questions of Van Wagenen may reasonably be testing something more than historical ability. For example one of the problems which is to test the pupil's ability in character judgment, describes the frontier dangers in the years, 1703-12, from Indian attacks and briefly sets forth the attitude

[12] Kelly, T. L., *Educational guidance: An experimental study in analysis and prediction of ability of high school pupils.* (Teachers College Contributions to Education, No. 71) New York: Teachers College, Columbia University, 1914, 116 pp.

of the frontiersmen in advancing the frontier despite persistent Indian opposition. The question is then put:

Draw a line under the three of the following words which you think best describe the action of these frontier people.

intrepid	dismayed	gentle	shrinking	resolute
cowardly	treacherous	timid	dauntless	deceitful

Does the average high-school pupil, one wonders, have a working vocabulary of such words as *intrepid* and *dauntless,* to select only two? In other lists from which a selection is to be made one finds such words as *irresolute, compliant, ignoble, undaunted, unscrupulous, wary, bigoted,* etc. How is a pupil to exercise his powers of character judgment, if we ask him to use tools with which in all probability he is unfamiliar? In many cases, is it not reasonable to assume that a simplification of such words in these particular tests might produce better results historically?

Of course one can always find questionable details in tests of information prepared by others. Here the reader is reminded, as stated above, that the question of the personal judgment of A is set over against that of B. As one studies the Starch test, one's sensibilities may be disturbed rather frequently by the appearance of annoying detail, as for example: New Hampshire was founded in . . . . To one not interested in the teaching of historical information per se, the recurring questions concerning military history seem rather far removed from the advisability and desirability of making history real to the pupil. The writer, however, has little doubt that considerable space was given to military history in the texts upon which this test is based, and has less doubt that not a small number of teachers are still emphasizing details of military history. The Starch test cannot be dismissed without saying a word concerning its rather poor arrangement. Sixty-nine questions are arranged indiscriminately, without any effort apparently at placing them into such groups as men-events, events-men, dates-events, etc. The psychological effect upon the subject of such grouping needs no emphasis. The Rayner test and the Hahn scale, with respect to

questions of detail, share in common to some extent with the peculiarities of the Starch test.[13] The Davis test, on this same score, in the opinion of the writer, has been rather unfairly criticized recently.

Approaching the tests in a spirit of constructive criticism, one is impressed by the method used by Van Wagenen for the testing of chronological sense. In the Informational Series there are questions of the following type, asking for the arrangement of the events in the order in which they occurred, the order to be indicated by numerals:

..... Purchase of the Philippine Islands
..... Missouri Compromise
..... Introduction of secret voting or the Australian ballot system in the United States
..... Annexation of Texas
..... Building of Panama Canal
..... Passage of the first protective Tariff Act
..... Secession of the Southern states

Space will permit but one example of the type of question found in the Van Wagenen Thought Scales:

Although agreement of peace was signed by the commissioners of both Great Britain and the United States at the city of Ghent in the Netherlands on Christmas Eve, 1814, the news did not reach America until after the Battle of New Orleans had been won by the Americans on January 8, 1815, with a loss of nearly 2,000 soldiers to the British.

(a) Why do you think the news was so long in getting to America?

(b) What effect would this victory of the American army have upon the arrangements for peace?

An example of the type of question found in the Van Wagenen Character Judgment Scale was given above.

Two exercises of the Harlan test for the testing of a knowledge of chronology are here given for what they may be worth. The year is asked for in the following events:

1. Discovery of America.....
2. Signing of the Declaration of Independence.....
3. Settlement of Jamestown, Va.....
4. Settlement of Plymouth Colony.....
5. First battle of American Revolution.....

[13] For Hahn's defense, see *Historical Outlook*, 13:12, January, 1922.

The most important event is asked for the following dates:

1. 1803.....
2. 1820.....
3. 1812.....
4. 1765.....
5. 1850.....

The Harlan practice of furnishing a list from which to make a selection might wisely have been followed in the above two exercises.

The Spokane test, in addition to its valuable feature already mentioned, contains an exclusive group of exercises requiring a working knowledge of civic terms. The Sackett United States History Scale contains a group of exercises classified as historical terms. The Bell and McCollum test and Sackett scales are the only ones providing for a map study.

The following is an example of the type of question involving judgment used by Rugg:

The English colonial policy succeeded while the French policy failed because

1. The English were better farmers than the French.
2. The climate in the English colonies was better.
3. The natural resources were more suitable.
4. The English pursued a policy of permanent settlement.
5. The English government at home was better.

The Barr Diagnostic Tests remain to receive brief comment. There is something very ingenious about these tests. Their strength grows on one. Unlike any other test in the field of history, they provide for a practice test. The resulting effect of a working acquaintance with the problem at hand can be readily appreciated. The type of question to be found in these tests is in the form of an average-sized paragraph setting forth, for example, a brief of the Dred Scott case which is followed by questions involving selection and some judgment. Several groups of the questions are:

(a) Put a cross (x) before each of the following statements that is discussed in the above paragraph:

1. The repeal of the United States Constitution.
2. The Dred Scott decision.
3. The repeal of the Missouri Compromise.

(c) The decision said that

1. Dred Scott was a citizen of the United States.
2. Dred Scott could not bring suit before the court.
3. Dred Scott was not a negro.

The chronological questions are similar to the Van Wagenen example given above, but provide many variations so that they furnish by far the best test for the testing of chronological sense. One can not resist including here one at least of the chronological questions:

Indicate approximately how long ago the following events occurred by placing one of the following terms before each: decade, two decades, century, one-half century, one-quarter century, two centuries, etc. Do not answer by giving the date.

(a) Declaration of the Monroe Doctrine
(b) The annexation of Texas
(c) Bacon's Rebellion
(d) The completion of the Panama Canal
(e) The battle of Lexington

General Criticisms

Little comment need be made upon the weak criticism that a standardized test can be used but once with a given group of pupils; no one denies the fact, and, of course, the same criticism holds true of any test, whether or not it falls within the classification of scientifically prepared tests. It is to be noted, however, that a number of the tests have more than one series. Of this type are the Barr, Davis, Buckingham, and Van Wagenen scales. The number of series are indicated in Table I. The available Starch test is of Series A; further series (B, C, and D) have been prepared, but have not as yet been published.

The preparation of a test acceptable to the greatest number, however, has not as yet been possible because of the decided lack of uniformity of opinion among teachers of history as to what constitutes content, aims, values, etc. It is easily seen, of course, that the social sciences do not lend themselves to objective testing in exactly the same way as does mathematics, for example. And it is when one prepares a test which aims at testing abilities in history other than the informational, that one puts a rod in

pickle for oneself. Sooner or later a key must be constructed for the scoring of the test. The line of demarcation between objective and subjective scoring in history is neither clear nor straight. In general the scoring of the tests—especially when one gets out of the informational field—is less objective than in most fields, and is generally neither simple nor accurate.

The details of scoring will not concern us much here. It is sufficient to note that the two scales of Sackett, the tests of Boston, Davis, Rayner, Rugg, Spokane, and Starch are wholly objective in scoring. The Barr test is on the whole objective, but a rather complicated weighting system is used in the scoring of parts of the test. The Harlan test and Hahn scale make provision in scoring for full and partial credit. The Van Wagenen Information Scales and Thought Scales provide for full, partial, and no credit within arbitrary limits. The Character Judgment Scale is scored objectively. The objective scoring is made possible in this particular scale by the utilization of the checking system. In the scoring of the Van Wagenen Scales, which test subjective material largely, subjective scoring is reduced to a low possible minimum. The objective system of scoring, however, must of necessity have rather arbitrary limits; the result is a cumbersome method of objective scoring. The scoring of the Kelly test is very complicated, so complicated in fact, as to make the test prohibitive for ordinary use, even though one would care to overlook its numerous other peculiarities.

A valid criticism of the present history test is the necessity of postponing administration of the test until the close of the year. Obviously if the test is to serve the primary purpose of furnishing the teacher with an appraisal of his own efforts at presentation of a subject, administration of the test must come at an earlier period. The solution of the problem in the field of informational history is simple. The Davis test, which represents exercises for the colonial period alone, has given us the simple solution. The determination of periods to follow the colonial period would not be a difficult matter; the outstanding textbooks agree in such broad divisions of the subject. It is

possible and convenient, however, to use some of the other available tests before the close of the year. For instance, the writer, for purposes of his own, has found it possible to use Van Wagenen's Thought Scale, Series B, together with Davis's Exercises at the mid-year examination period. The above criticism, moreover, can not apply to Van Wagenen's Character Judgment Scales. It holds, however, largely against all other tests in the field of history.

Whatever opinion one may share as to what constitutes a course in American history, all, no doubt, will agree that the Spanish-American War was not the close of history for this country. The slighting of the past two decades in American history on the part of the makers of tests is significant. But then it is a comparatively recent move on the part of textbook writers to "revise" their books so as to include the recent period. The last question, in point of time, found in the Rayner test is one concerning the campaign of 1896. Nor does the Starch test reach the present century for historical content. The answer to one question in the Harlan test might be made on the basis of historical events subsequent to the Spanish War; it might, on the other hand, be answered as accurately by recalling events prior to the War. There is in the Harlan test one question, and only one, which definitely tests the current period. However desirable the Barr Diagnostic Test may be in general, it is weak in its treatment of the recent period. The Spokane test, on the other hand, devotes fully 20 percent of its questions to the past two decades. The Van Wagenen Information Scale, Series A, devotes a fair proportion of its questions to the recent period; Series B, to the contrary, with one notable exception ignores American history since 1900.

General Value of Tests

Distinctive values of individual tests have already been noted. The general value of the tests will be given in summary statements, resort being made in some cases to the opinions of others. Probably few teachers of Roman history present material in the

detailed fashion in which the Barnard test would attempt to test information in that period. Wilson and Hoke point out in regard to the Sackett Ancient History Scale that "its general acceptance would be detrimental to the proper teaching of history."[14] Sackett himself recognizes the limitations of the scale in that it "measures only a limited and specific, though not an unimportant, kind of historical ability."[15] For diagnostic purposes there is nothing better in the field of American history than the Barr Diagnostic Tests. It was the purpose of the author that the test be primarily diagnostic and only secondarily was it intended to measure achievement in history. The Davis test is the only one which confines itself to a particular period of information, thus producing a test which can be administered during the early part of the course. Forty-four different tasks in 12 minutes, however, suggests a test giving little time for recall of facts even though wholly mechanically learned. Both the Davis and Barr tests, as was pointed out above, lack standardization. The Harlan Information Test is, in the opinion of Rugg, probably the best of the completion tests. The Rayner test, not at present available, has, in the opinion of the writer, no particular distinction as an information test. One of the most significant values of the Sackett United States History Scale is its accuracy in grading. The Spokane authorities have informed the writer that the Spokane test is there used only in the grades. It is a test, however, that might well be used in secondary schools, if it were standardized. One excellent feature of the Spokane test—the inclusion of civic terms—has been noted. The Starch test like the Rayner test permits of unlimited time for administration. The Bell and McCollum test and study is valuable particularly for its results. The Buckingham test was the pioneer in the use of thought. The manner in which the Kelly test is weakened in its prognosis of historical ability needs no further comment. The complicated scoring system would scarcely

[14] Wilson, G. M., and Hoke, K. T.: *How to measure*, New York: Macmillan Company, 1920. pp. 225-26.

[15] Sackett, L. W., "Scale in ancient history," *Journal of Educational Psychology*, 8:285-93, May, 1917.

recommend its use, if all other shortcomings of the test were ignored. The Rugg test, apparently not available at present, is a move in the right direction; its value is limited, however, until standards are established.

It is the writer's opinion that the Van Wagenen scales mark the highest point yet reached in tests for the field of history, and that they point out the general direction which future tests will probably wisely follow. Scientifically founded and carefully standardized, these tests may well stand on their own merits. Separate scales for information, thought, and character judgment, on the part of Van Wagenen, have begun a movement which might profitably be carried on by future makers of tests in the interests of both teachers and pupils. References have already been made to the outstanding merits of these scales. One additional merit should be noted. A weighting system is used in scoring which gives greater values to questions which are located nearer the medians of grade distributions.

Summary

This brief survey of history tests has attempted to present: (1) the methodology of constructing tests; (2) the types of tests in the field and their nature; and (3) the criticism, destructive and constructive, passed upon the various tests. With regard to the last point, it will be agreed that many of the criticisms have been surface criticisms. The test movement in general is, of course, no longer on the defensive; and it is fair to assume that history tests in particular are now approaching, rather slowly perhaps, a clearly defined goal. In history tests, and in fact, in the whole test movement as in our entire educational system, there is both room and need for improvement.

ANOTHER CRITICISM OF TESTS REQUIRING ALTERNATIVE RESPONSES

C. W. ODELL
University of Illinois

In a recent issue of the JOURNAL OF EDUCATIONAL RESEARCH there appeared an article entitled "A Criticism of Tests Requiring Alternative Responses" in which the criticism offered was almost entirely adverse. It is the purpose of the writer to discuss briefly certain of the objections to this form of test contained in that article and to present certain arguments for the use of such tests.

One of the chief criticisms of this form of test is that the common practice of subtracting the wrong answers from the right ones does not eliminate the effect of guessing. It is stated that if one uses this procedure he must assume that individuals in such tests always guess an even number of times and further that if an individual happens to guess wrong the first time his second guess must be right; his third, wrong; his fourth, right; and so on. As a matter of fact neither of these assumptions is involved in the method of scoring referred to. The writer does not know of any advocate of these tests who claims that subtracting the wrongs from the rights gives the true score for every individual taking the test. Those who believe in using these tests do, however, claim that this method of scoring will more often give the true score than will taking the number of rights as the score. It can be proven mathematically that in such an experiment as was described in the article referred to, drawing at random from a sack containing the same number of white and black buttons, if the number of drawings made is even it is more likely that half of the buttons taken out will be black and the other half white than that any other combination will result. If the number drawn out is odd it is most likely that the number of one color will be just one greater than the number of the other color. It is true, as was suggested, that the activity of the human

mind is not as mechanical a process as drawing buttons out of a sack. If, however, as should be the case, the number of exercises requiring positive answers is practically the same as the number requiring negative answers and if the students know this fact, the distribution of their responses will tend to be the same as that of buttons drawn out of a sack under the conditions named above. This is stated not merely upon theoretical considerations but is also based upon certain data collected by the writer and others at various times.

The second assumption referred to above is also erroneous. Even if we assume that the number of right guesses is always the same as the number of wrong guesses, which we do not, it is not necessary that they should be made in exactly alternate order. There may be two right guesses followed by two wrong guesses, etc, or one right guess, three wrong, two right, etc., or any other arrangement so that half are right and half wrong.

In the article referred to a table is presented which shows the results of the drawing experiment described therein. The different series of drawings are each supposed to represent one student's responses to an alternative test and upon that basis the scores to be assigned are discussed. The writer has discussed the method of scoring given with several persons who are in the habit of using this type of test and has not yet found one who would score the data given as was done in the discussion. Therefore he does not believe that the discussion based upon that method of scoring is a valid argument against this type of test.

It is further stated that to justify subtracting the wrong answers from the right the assumption must be made that every wrong answer is a guess. This, however, is not the case. Suppose, for example, that a student has twenty correct and five incorrect answers to a list of twenty-five exercises. The five incorrect answers may be divided in any ratio between pure guesses and answers which the student incorrectly believes are right. It is a rather generally accepted maxim among educators that when we know something we should know that we know it and, furthermore, that it is better to know that we do not know

something than to think we know it when we do not. Therefore if the student in question thinks that he knows the correct answers to all the exercises but really gives the incorrect answers to five of them a deduction should be made from his positive score of twenty. In other words, according to this system of scoring a student who knows twenty and does not attempt the other five will receive a higher score than a student who thinks he knows all of them but is mistaken in some cases. This is as it should be.

It is further asserted that one must assume that because a certain number of answers are wrong the same number of right answers must be guesses. As was stated above in discussing the first objection to this type of test this assumption is not necessary. The assumption made is that if the wrong answers are all guesses it is more likely that the same number of right answers will be guesses than that any other condition will exist. In either case, whether the wrong answers are all guesses or whether some of them are given in the belief that they are correct, the true score is most nearly arrived at by subtracting the number wrong from the number right.

In considering the justification for subtracting the wrongs from the rights let us also consider the cases of two other students. Student A, who is of the confident type, knows the answers to fifteen out of twenty-five exercises and guesses at the other ten. Student B knows the subject as well as Student A but is of the timid, hesitant type. He answers fifteen exercises correctly but does not make guesses upon the other ten. Does Student A deserve a higher score than Student B merely because his mental disposition results in his guessing, whereas Student B's prevents him from guessing? Most persons will, I believe, answer this question in the negative. Therefore, to allow for the exercises Student A happens to guess correctly, it is necessary to subtract the wrongs from the rights.

The writer does not wish to close the discussion of this method of scoring without mentioning the fact that in case each student answers all the exercises it really makes no difference

whether the scores are given according to the number right or according to the rights minus the wrongs. The absolute scores, of course, do not agree, but the relative position of each individual in the class distribution remains the same. These tests are, however, usually given as rate tests, that is, they have a time limit so short that all students cannot complete them. In view of this fact it is necessary to subtract the wrongs from the rights to approximate the true scores.

The article referred to further attacks alternative response tests upon the ground that they encourage negative and dependent achievement. The writer certainly admits that these types of achievement do not have the value that positive independent achievement has but he does not admit that they are entirely without value. Not infrequently stating what a thing is not is of great assistance in defining what it is. It is sometimes more important that the student know that two things are not the same than that he be able to give a definition of both. Dependent achievement also has its place. Outside of school work it is often desirable to identify correct answers as well as to produce them. Therefore it seems illogical to exclude this type of achievement from our school work.

Several other objections to such tests are given. It is said that they cannot be used for diagnostic purposes and that they encourage reactionary ideals of instruction. The writer admits that these objections have some merit, especially if no other type of test is used. In using this type, however, it is his belief that others should be used as well.

Tests of this sort have certain positive advantages that render their use of great value in teaching. By using them it is possible to cover a much wider range than can be covered by the usual form of test within the same time. Instead of having the students spend ten minutes discussing some one point, 25 or more different facts can be tested in this time. It is true that they do not offer much opportunity for telling why, but it is not necessary that such opportunity be offered in all tests. We often desire to measure actual knowledge that students have acquired.

Another advantage of these tests is that they measure the subject that it is desired to measure. The ordinary written test is often as much a measure of speed of writing and of freedom from writer's cramp as it is of knowledge of the subject supposed to be tested. The alternative test reduces the element of writing to a minimum so that it has practically no effect upon the amount of work done.

Finally the ease and objectivity of scoring which this type of test affords gives it decided advantages. Anyone who has taught will appreciate the labor involved in scoring a set of papers of the ordinary discussion type. Even after devoting considerable time and energy to the job the instructor usually feels doubtful as to how well it is done. On the other hand, alternative response tests can be scored very rapidly, those of 25 and 30 exercises not requiring as much as a minute apiece. The scoring is also absolutely objective. This latter advantage alone is sufficient to offset more or less weighty faults. Elsewhere in the same number of the JOURNAL OF EDUCATIONAL RESEARCH that contained the criticism of these tests is an excellent example of the lack of objectivity often found in scoring test papers. A fourth-grade geography and an eighth-grade arithmetic paper were scored by 92 teachers, the scores upon the former running from 0 to 74 percent and those upon the latter from 10 to 100 percent. Many other cases could be cited which show similiar variations. It is evident that in such cases the score which one particular teacher happens to assign has a very doubtful meaning. Certainly a test that yields an absolutely objective score is preferable *on that ground alone* to almost anything upon which the scores have practically no meaning.

A MENTAL SURVEY OF FIFTY-NINE NORMAL SCHOOL STUDENTS: SOME CORRELATIONS AND CRITICISMS

LAWRENCE A. AVERILL
State Normal School, Worcester, Massachusetts

The data for this study were derived from the following four sources: (a) the application of the Otis Group Examination Scale to 59 entrants to normal school, comprising the entire two-year group entering in September, 1920; (b) the application of the Thurstone Psychological Examination to the same group; (c) the analysis of scholastic grades awarded to the 59 students in the first half-year of their normal-school course; and (d) instructors' estimates of their native ability, as judged by their classroom and general school behavior.

The Otis test.—The scale used in this study was form A, of the 1919 edition. The results may be summarized as follows:

Range, 206-119; median, 159; average, 162; Q_1, 147; Q_3, 177; quartile deviation, 15; mean deviation, 16.5.

The Thurstone test.—The scale used was Test IV for college freshmen and high-school seniors. The results may be summarized as follows:

Range, 147-50; median, 77; average, 82; Q_1, 71; Q_3, 92; quartile deviation, 10.5; mean deviation, 12.7.

School grades.—In order to reduce to a numerical basis the grades awarded to the 59 students during the first half-year of their attendance, a number was substituted for each of the five letters employed: 1 for A, 2 for B, 3 for C, 4 for D, and 5 for E. All grades received (10 in number for each individual) were averaged numerically.

The results of this process for the 59 students in question may be summarized as follows:

Range, 1.4-3.7; median, 2.3; average, 2.3; Q_1, 2.6; Q_3, 2.0; quartile deviation, 0.3; mean deviation, 2.4.

Instructors' estimates.—For the purpose of securing the data desired in this connection, those five instructors who were best acquainted with the work of the students and who would thus presumably be in the best position to judge them on the basis of general intelligence, were asked to indicate a numerical grade for each student. They followed the same scheme as was employed in obtaining averages for classroom grades; i.e., a student whose native ability was judged by the instructor to be of "excellent," or "A" quality, was ranked 1, while any student whose intelligence might be adjudged to be of the "failure" or "E" quality, was ranked 5. All intermediate judgments between "excellent" and "failure" were recorded by appropriate numbers.

The results of averaging the five grades thus derived may be summarized as follows:

Range, 1.0-3.6; median, 2.0; average, 1.9; Q_1, 2.2; Q_3, 1.4; quartile deviation, 0.4; mean deviation, 0.5.

TABLE I. A CONSPECTUS OF THE DATA DERIVED FROM THE FOUR SOURCES

Source	Range	Median	Average	Quartile Deviation	Mean Deviation
Otis Test	206—119	160.0	162.0	15.0	15.5
Thurstone	147— 50	77.0	82.0	10.5	12.7
Grades	1.4—3.7	2.3	2.3	0.3	2.4
Instructors' Estimates	1.0—3.6	2.0	1.9	0.4	0.5

TABLE II. FREQUENCY DISTRIBUTION: OTIS AND THURSTONE SCORES

Otis Test		Thurstone Test	
Scores	Frequency	Scores	Frequency
115—126	2	45— 56	2
127—138	7	57— 68	9
139—150	7	69— 80	24
151—162	16	81— 92	11
163—174	10	93—104	8
175—186	10	105—116	3
187—198	5	117—128	1
199—210	2	129—140	0
		141—152	1
Total	59	Total	59

It is apparent from the results that the Thurstone test reveals much greater ranges in intelligence than does the Otis, the highest score being only approximately 75 percent higher than the lowest score in the Otis test, while in the Thurstone test the highest score rises about 200 percent above the lowest. In other words, it would seem that the Thurstone test diagnoses either the points of extreme weakness, or of strength, or of both, in the individuals tested to a greater degree than does the Otis.

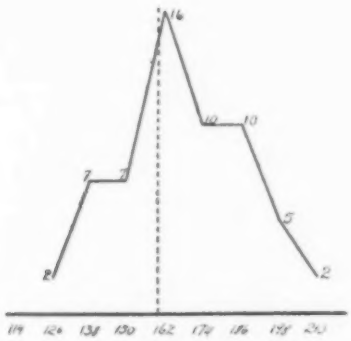

FIGURE 1. FREQUENCY DISTRIBUTIONS FOR THE OTIS TEST

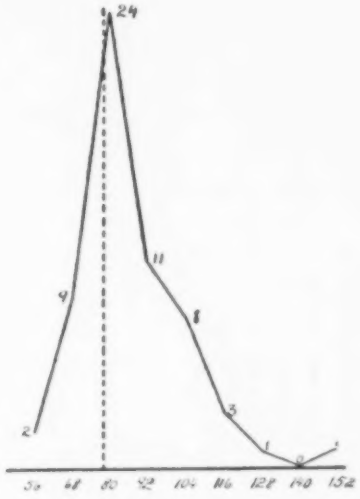

FIGURE 2. FREQUENCY DISTRIBUTIONS FOR THE THURSTONE TEST

Table II and Figures 1 and 2 show that the median for the Otis test is much nearer to the middle of the range of scores than is the case with the median of the Thurstone test. An analysis of the type of questions used in the Otis and Thurstone tests reveals a possible explanation of this relative proximity of the median score to the middle of the Otis series and of its relative proximity to the lowest score in the Thurstone series. The Otis test contains a large number of problems requiring keenness of language perception, mathematical judgment, logical perception, etc. The

Thurstone test contains, in addition to these, a new type of question; namely, the *general information* question. During the scoring process, the number of errors made was entered under five headings: *information, word relation, logical relation, number relation,* and *comparison.* The average percent of errors made by the entire group of 59 individuals was then computed. It was found that 40 percent of them (39.7 percent) were errors in answering the questions of general information. Inasmuch as this type of question does not appear in the Otis test, while the others do, it seems logical to conclude that the reason for the relatively low median in the Thurstone group is the prominence in that test of questions of general information, inability to answer which caused 40 percent of the failures. The remaining 60 percent of erroneous answers, while I did not work out percents for them, appear by inspection to be tolerably evenly distributed among the four other types of questions. This would indicate that errors in the general information questions are nearly 300 percent more numerous than under any other caption, which condition would, as stated before, be a large factor in explaining the much lower grouping of scores in the Thurstone as compared with the Otis test.

The range of scholastic abilities of the students, as represented in the grades earned by them during the first half-year's work, is from 1.4 to 3.7, or a range of nearly 200 percent above the minimum grade. This would indicate that the Thurstone test, in which a similar range of 200 percent occurs, is a better discoverer of the extremes of the class abilities than is the Otis, in which the range is 75 percent only. The median grade and the average grade, however, are more nearly suggestive of the Otis results, inasmuch as they are at the approximate midpoint of the scores.

The instructors' estimates of general intelligence cover a still wider range; that is, 1.0 to 3.6, or a difference of more than 250 percent between lowest and highest. The median and average grades, while slightly above the midpoint, are here again suggestive of the Otis results. The conviction then is that the

Thurstone test discovers the individual limits better, while the Otis test agrees more accurately with the scholastic grades and instructors' estimates of the students as a mass.

Correlations.—The correlations shown in Table III were worked out by the Pearson formula.

TABLE III. CORRELATION COEFFICIENTS

Otis—Thurstone	0.62
Otis—Instructors' estimate	0.64
Thurstone—Instructors' estimate	0.36
Otis—Average school marks	0.61
Thurstone—Average school marks	0.38
Instructors' estimates—Average marks	0.68

Rearranging the coefficients of Table III, for greater ease of comparison, we have:

	Correlation with Instructors' Estimate	With Scholastic Attainment
Otis test	0.64	0.61
Thurstone test	0.36	0.38

The Otis and Thurstone tests, while correlating highly with each other, fail to correlate similarly high with the instructors' estimates of ability and with the school grades received. As we should expect from the closer agreement of the Otis scores with the abilities and grades of the students as a mass, the correlation between the Otis test and the instructors' estimate of general intelligence on the one hand, and with the scholastic attainments of the students on the other, is very appreciably higher than is the correlation between the Thurstone test and those factors, the Otis agreeing with the estimated general intelligence and scholastic achievement of the group approximately 75 percent more closely than the Thurstone results. One of two interpretations of this condition is obvious: either the school standing and school estimate of the students fail conspicuously to do them full justice from the point of view of their real worth as discovered by the Thurstone test, or else, granting the school grading and estimate of the group to be reasonably accurate, the Otis test

would seem to be a much surer means of determining their intellectual status than the Thurstone. It is difficult to determine with any certainty which of these possibilities is the more probable. If, as stated before, the reason for the Thurstone test results being grouped so generally toward the lowest individual scores is the prominence of questions of general information in that test, and the disproportionately large number of failures in answering correctly such questions of general information, it would seem a fair assumption that, inasmuch as the basis of marking and estimating the intelligence of normal school students is largely concerned with personality, teaching promise, mastery of elementary subject matter, aptness and readiness in class discussions, etc.—with little opportunity being offered for judging a student in terms of general information possessed—the explanation of the relatively high coefficient of correlation of the Otis test with scholastic attainment and instructors' estimate would lie in the absence in that test of questions involving general informaion. For the same reason, it would seem a possible assumption that the Thurstone test, judged on this sole point, is a more accurate gauge of general intelligence than is the Otis.

If, on the other hand, such knowledge as tests involving general information call for is not a proper index of general intelligence, then the possibility is that the school standing and estimate are properly arrived at and the Otis test may be a better discoverer of general intelligence than the Thurstone. The sort of information which must be possessed by one who is to score high in the Thurstone test is whether Blanche Sweet is a writer, singer, suffragist, or movie actress; whether shoes are made by Swift, Smith and Wesson, Babbitt, or Douglas; whether Yale University is at Annapolis, Ithaca, Cambridge, or New Haven. To succeed well on the general information part of the test one must know what a Corona is; what the Delco system signifies; the identity of Irvin Cobb, of Blackstone, and of Kelvin. The decision as to the importance of this type of general information in estimating intelligence is at present dependent mainly on opinion.

Finally, as was to be expected, the highest correlation of all was found to be between the estimate of general intelligence and the scholastic standing of the students: 0.68 (Pearson). The correlation is almost exactly the same as that existing between the Otis test and these same factors of instructors' estimate and scholastic achievement, the coefficients being 0.64 and 0.61, respectively.

TRANSMUTATION OF SCORES BETWEEN BINET TESTS AND GROUP TESTS

W. T. Root
University of Pittsburgh

Recently the writer published some correlations between Binet Tests and certain Group Tests.[1] The tabulations presented here are derived from the correlation plottings, the scores being transmuted by a method of rank correspondence.[2] For this purpose, all of the grades involved in a particular group test have been thrown together.

As indicated by Table V, the transmutation of Binet scores with Haggerty Delta I, grade by grade, the mass scores for a particular mental age tend to be increasingly higher with the increase in school grade. Consequently, in the mass test, one may consider the score at the school grade (most likely to contain a child of the age in question) as most nearly comparable to the corresponding Binet score. A grade below the median will tend to yield a somewhat lower mass test score for a designated Binet mental age; a grade above the median grade a somewhat higher score.

Otis' assumption of correspondence between rank-orders is at times far from true even as a rough approximation. When a row in a scatter-distribution nearly runs the gamut of the columns, and when a column similarly cuts the rows, one is naturally skeptical of the transmuted scores. However, for rough cross-comparison, the accompanying transmutations should be of value for administrative purposes. The correlation coefficients given in detail in the previous article[1] indicate somewhat the reliability of the transmutations.

---

[1] Root, W. T. "Correlations between Binet tests and group tests," *Journal of Educational Psychology*, 13:286-92, May, 1922.

[2] Otis, A. S., and Knollin, H. E. "The reliability of the Binet scale and of pedagogical scales," *Journal of Educational Research*, 4:121, September, 1921.

TABLE I. MASS TESTS TRANSMUTED INTO BINET SCORES

Binet (age in years)	Otis Primary A	Otis Advanced A	Haggerty Σ I	Haggerty Δ I	Haggerty Δ II	National A.I	National B. I	Terman Group A	Mentimeter	Dearborn Series I	Dearborn Series II	Illinois I	Illinois II
3									0				
3½	1		0	0					0	5			
4	5		1	1					1	14			
4½	8		1	2					2	20			
5	12		1	3		4	4		3	38			
5½	14		1	9	7	7	8		4	57			
6	20		2	13	8	10	14		5	77			
6½	27		3	19	10	15	20		11	93			
7	33		4	25	13	22	26		18	115		17	
7½	38		5	32	17	30	32		23	128	23	23	
8	43		9	39	22	39	41		27	148	30	27	
8½	48		16	48	32	47	49		38	168	35	34	
9	54		25	55	42	59	56		49	200	42	41	
9½	58	38	30	60	48	70	64	53	58	217	49	45	
10	62	53	32	63	54	77	71	57	65	224	56	51	68
10½	64	64	34	67	58	83	79	62	71	230	64	55	70
11	66	72	36	70	67	90	89	67	74	235	69	61	73
11½	68	81	38	74	75	99	95	72	78	239	78	70	79
12	69	88	40	78	82	104	103	78	83	243	85	77	83
12½	71	94	41	83	88	113	113	83	86	248	90	85	87
13	74	101	42	87	96	121	120	90	89	253	98	90	93
13½	78	107	44	96	103	127	126	96	92	258	105	95	98
14	80	113			110	134	130	102	94		113	101	103
14½		119			115	138	136	109	93		120		107
15		124			118	143	139	114	100		127		109
15½		130			123	149	145	120	103		134		113
16		136			128	154	151	126	107		141		118
16½		142			131	159	156	131	110		148		123
17		147			136	164	159	139	114		155		131
17½		153			140	167	163	148	118		165		136
18		162			143	169	166	157	122		174		142
18½		168			146	171	171	174	131		185		147
19		179			150			186	142				
19½		194						198	159				
No. cases	198	218	160	162	232	211	210	160	407	156	195	100	107
G'des	1–4	5–12	1–3	1–3	3–9	3–8	3–8	7–12	1–12	1–3	4–9	3–5	6–8

TABLE II. MASS TESTS TRANSMUTED INTO HAGGERTY SCORES

Haggerty ΣI with:			Haggerty ΔII with:					Haggerty ΔI with:		
HΣI	O.P.A.[a]	D.S.I.	HΔII	O.A.A.	N.B.I.	Ment.	D.S.II.	HΔI	D.S.I.	Ment.
2	4–26	9–104	5					2	32	2
4	33	120	10		(19)			5	52	5
6	38	140	15		28		26	8	62	8
8	39	151	20		36	36	28	11	70	10
10	41	156	25		42	39	31	14	85	12
12	42	162	30		47	42	33	17	100	14
14	44	167	35		50	46	35	20	108	16
16	46	172	40		56	50	38	23	112	17
18	49	176	45	32	62	55	44	26	117	19
20	50	187	50	40	68	60	49	29	123	20
22	52	192	55	47	70	65	55	32	130	22
24	53	199	60	52	78	69	60	35	139	25
26	54	204	65	58	86	72	63	38	152	28
28	55	208	70	61	92	74	65	41	160	31
30	56	211	75	67	98	76	68	44	163	33
32	58	219	80	75	104	78	73	47	166	35
34	61	226	85	84	108	82	79	50	171	37
36	64	231	90	92	116	85	86	53	176	39
38	68	240	95	99	121	89	96	56	187	41
40	72	248	100	104	126	92	104	59	200	45
42		254	105	109	130	94	109	62	216	54
			110	116	134	97	113	65	220	60
			115	122	138	99	118	68	224	62
			120	130	142	103	124	71	232	66
			125	137	147	107	133	74	240	72
			130	141	152	111	146	77	248	81
			135	152	158	116	159			
			140	165	162	119	168			
			145	184	164	122	173			
			150			127	178			
No. cases	160	155		154	206	192	172		118	153
G'des	1–3	1–3		5–9	3–8	4–9	4–8		1–3	1–3

[a] For the meaning of the abbreviations at the top of these columns see Table I.

TABLE III. MASS TESTS TRANSMUTED INTO NATIONAL A-I SCORE

National A I	National B I	Otis Primary A	Otis Advanced A	Haggerty Δ II	Mentimeter	Dearborn Series II
9	9	15			10	
16	21	24		11	17	
23	28	31		15	24	22
30	37	37		19	30	23
37	44	42		25	35	28
44	49	45		31	40	34
51	53	48	24	36	45	37
58	58	51	28	43	50	40
65	65	55	37	48	56	46
72	70	60	46	53	63	52
79	75	63	54	60	66	55
86	82	67	59	64	71	58
93	93	70	65	70	73	64
100	100	72	72	79	78	69
107	107	74	79	85	82	76
114	114	76	90	89	86	85
121	121		100	99	91	96
128	128		109	107	96	110
135	135		119	112	99	119
142	140		127	119	101	127
149	145		134	124	105	137
156	153		143	130	108	148
163	160		155	139	118	173
170	170		167	147	129	185
No. cases..	207	74	134	209	211	167
Grades....	3–8	3–4	5–8	3–8	3–8	4–8

TABLE IV. MASS TESTS TRANSMUTED INTO TERMAN GROUP SCORES

Terman group test A	National A I	National B I	Otis Advanced A	Haggerty Δ II	Mentimeter
55					65
60	98	102	76		70
65	106	107	80	75	77
70	112	112	84	92	81
75	118	118	89	94	83
80	120	122	95	96	85
85	124	126	99	98	87
90	127	128	102	101	89
95	130	130	106	104	91
100	132	132	111	109	93
105	135	134	115	113	96
110	139	137	119	115	98
115	143	139	121	117	99
120	145	141	127	120	103
125	151	144	133	125	106
130	156	148	138	127	110
135	159	152	141	131	113
140	160	156	146	133	114
145	161	158	150	135	115
150	162	159	153	137	117
155	163	160	157	138	119
160	164	161	161	139	121
165	165	162	165	144	124
170	167	166	168	147	126
175			171	149	128
180			173		130
185			177		136
190			184		144
195			189		151
200			195		
No. cases....	76	76	160	97	159
Grades.....	7–8	7–8	7–12	7–8–9	7–12

TABLE V. TRANSMUTATION OF SCORES—BINET WITH HAGGERTY-DELTA I, GRADE BY GRADE

Binet	Haggerty Δ I			
	All	Grade I	Grade II	Grade III
3				
3½	2	2		
4				
4½				
5	3	3		
5½	9	9	14	
6	13	11	17	
6½	19	15	23	27
7	25	23	25	32
7½	32	33	35	38
8	39	38	41	43
8½	48	46	53	50
9	55	57	59	65
9½	60		65	60
10	63			63
10½	67			67
11	70			70
11½	74			76
12	78			80
12½	83			85
13	87			87
13½	96			96
No. cases..	161	88	36	37

Editorials

The reader's attention is called to the note on page 349 headed "Erratum."

PARTIAL CORRELATION

We have desired for some time to say something editorially concerning partial correlation. The reader is permitted at this point to reflect upon the natural depravity of a soul longing to express itself on such a subject. The writer has many friends whose aspirations are quite otherwise. Some of them get no further than aspirations, fatuously relying upon the sufficiency of mere elevation of aim. Moreover, they will have nothing to do with other aims. They resemble the old lady who said she didn't like spinach and added that she was glad of it; because if she did like it she'd eat it, and she couldn't bear it. Now, if you whose eye has fallen on this page, resemble the old lady; and if things statistical are to you abhorrent, partake not lest you acquire a liking for something you detest.

Having thus assured ourselves of congenial company from this point on, let us proceed. The term "partial correlation" is well chosen. It means the degree of association between two things after the influence of other things is taken away. It is one of the ways in which the idea "other things being equal" is put into effect. Suppose we wish to investigate the relation of the salaries of teachers to their teaching ability—indeed, we urgently suggest this as a subject of inquiry. We ought to know to what extent by paying more dollars to teachers a community gets better teaching.

Assuming the existence of a good measure of teaching ability, we may find for each of a large number of teachers distributed throughout the country two things, namely, salary and success in service. Let us call these two things variables for the sufficient reason that both of them vary among different individuals. If, as is presumably the case, they tend to vary together—that is, if

in general high pay goes with good service and vice versa—then the relationship between these two variables is direct. The strength, however, of this tendency to vary together may be strong or weak. This means that few or many exceptions may occur—exceptions in virtue of which a poor teacher may receive high pay or a good teacher, low pay. Evidently we need a numerical expression of the strength of the tendency of these variables to rise and fall together.

The correlation coefficient is such an expression. Most readers of this JOURNAL are more or less acquainted with it. Let us suppose that for the variables, salary and teaching ability, it is found to be 0.5. This may mean that the tendency for salary and service to run together is "present and marked." Whether it has this meaning depends on the number of measures of salary and teaching ability from which it was derived and the degree of dispersion or "scatter" of these measures. If our cases are numerous and the measures do not greatly scatter, the coefficient 0.5 indicates fairly close relationship between the variables.

When, however, we try to read meaning into this expression of relationship, we are baffled. We are especially at sea, if we seek to base administrative action upon it. We cannot, for example, in the absence of additional data, infer causation—neither that good salaries produce good teaching, nor that good teaching ability commands good salaries. All we know is that, as conditions are, salaries and teaching ability tend to rise and fall together. Both, however, may depend upon a third fact or variable to such an extent that the real way for boards of education to secure good teachers may be to attend not to salaries but to this third variable. On the other hand, under the same conditions as to causation, the real way for a teacher to secure a better salary may be to reckon with this third variable rather than to make direct efforts to teach well.

In connection with salaries and teaching ability there are no doubt many other factors, each of which may exist in different amounts and each of which may vary directly or inversely with salaries and teaching ability. Such a system of variables has

been called a realm of concomitant variation. Within it our present statistical methods permit us to consider a few variables and to determine their effects. Here we come upon the method of partial correlation.

Our general knowledge of conditions, let us suppose, leads us to believe that one of the variables in closest correspondence with salaries and teaching ability is the amount of professional training a teacher has received. And suppose, further, that at the time referred to above when we obtained the salaries and teaching abilities of a large and representative group, we also secured for each teacher a numerical expression of his professional training. Finally, let it be granted that the relation between salaries and training gives us a correlation coefficient of 0.6 and that the relation between teaching ability and training yields a coefficient of 0.4.

Before looking into the implications of these data, it will be well to introduce the usual notation. As we all know, the symbol for the most acceptable correlation coefficient is r. In order that we may indicate in connection with r the two variables between which it expresses relationship, we often give code numbers to the variables and attach these code numbers to r as subscripts. If, then, we call salaries 1, teaching ability 2, and training 3, our suppositions in this realm of concomitant variation are: $r_{12}=0.5$, $r_{13}=0.6$, and $r_{23}=0.4$.

In investigating the real effect of salary upon teaching ability, we desire that other relevant things should be equal. If they are not, we are likely to be misled by a spurious relationship. One of the things which is quite evidently so relevant that it must be made "equal" is this matter of training. Partial correlation provides one way of doing this. It enables us to express the relation between salaries and teaching ability among those having the same amount of training. Evidently this result can be accomplished by a direct method. We may select from the original group of teachers all those who have the same amount of training and correlate salaries and teaching ability—i. e., find r_{12}—for this selected group. The resulting r-value is customarily writ-

ten $r_{12.3}$, the subscript to the right of the point indicating that the third variable (professional training) has been rendered constant, or "counted out." When a sufficiently large selected group can be secured, the direct method is the preferred procedure. It is not, however, the method of partial correlation, as that term is usually employed.

This method deduces $r_{12.3}$ from a formula in which the coefficients r_{12}, r_{13}, and r_{23} are used. The formula is

$$r_{12.3} = \frac{r_{12} - r_{13} \times r_{23}}{\sqrt{1 - r^2_{13}}\ \sqrt{1 - r^2_{23}}}$$

If we substitute the values 0.5, 0.6, and 0.4 for r_{12}, r_{13}, and r_{23}, respectively, $r_{12.3} = \frac{0.5 - 0.6 \times 0.4}{\sqrt{1 - (0.6)^2}\ \sqrt{1 - (0.4)^2}} = 0.35$. Thus it will be seen that the original relationship of 0.5 between salary and teaching ability reduces to 0.35 when the effect of training is eliminated. In other words, training contributed something to the apparent relationship between salary and ability. We cannot yet say that "other things" have been made equal, but we can say that one of them has been.

The formula above stated is worth studying. Note that the first term of the numerator is the gross coefficient for the same two variables as those which we are seeking to relate after rendering training constant. In other words, it is r_{12}. Note next that this is reduced by the product of the coefficients for each of these two variables and the third—i. e., r_{13} and r_{23}. If this product is greater than r_{12}, the resulting partial correlation is negative and the effect of rendering the third variable constant is to reveal an inverse relationship between the first and second variables. If, for example, the gross relation between salary and training and between ability and training were in each case 0.71 or more, the relation between salary and ability would not only be wiped out but actually reversed—this, of course, on the assumption that r_{12} remains 0.5. On the other hand, if the product of r_{13} and r_{23} is less than 0.5 (as it is in the case we have considered), then the resulting partial coefficient is positive.

To a certain extent, therefore, the numerator of the formula becomes a criterion of the character of the partial coefficient. If the product term is large with respect to the first term, the partial coefficient is small, or even negative; if it is small with respect to the first term, the coefficient is large.[1] The numerator is much more potent than the denominator in determining the partial coefficient.

Let us consider another variable. Experience is undoubtedly related to salary. It is probably, though certainly not so strongly, related to training; that is, there is probably a discoverable tendency for well-trained teachers to remain long in service. We may call experience, variable number 4. Let us suppose that for the large group of teachers we are considering, salary is related to experience to the extent of 0.7 (r_{14}), that ability is associated with experience to the extent of 0.3 (r_{24}), and that the correlation between training and experience is 0.1 (r_{34}).

When we attempt to make experience as well as training constant, we find that the relation between salary and ability is still further reduced. In fact it becomes only 0.23. This fact is expressed as follows: $r_{12.34} = 0.23$. Since two variables are rendered constant or of no effect, this expression is called in statistical parlance a partial coefficient of the second order. If the effect of one variable is counted out, as in the case of $r_{12.3}$, the result is called a partial coefficient of the first order. The usual coefficient (such as r_{12} or r_{23}) in which no variable is rendered ineffective is called a coefficient of order zero. It is not a partial coefficient. Evidently, then, the order of a partial coefficient depends upon the number of variables that are made constant. In the accepted notation this means the number of digits to the right of the point in the subscript.

Evidently, this process may be carried much further. More things thought to affect either salary or teaching ability may be made equal—such as age, intelligence, personality, type of position, the supply of teachers in relation to the demand for them, the purchasing power of the dollar, the standard of living, the

[1] True only if the gross measures of relationship are all positive.

conditions of service. By securing quantitative measures of these and by applying the method of partial correlation in analyzing the most important of them, we may arrive at a very much more precise knowledge of the extent to which salary really depends upon ability and of the extent to which it depends upon other things related either to it or to ability.

The method has in the past been distressingly tedious. Lately, however, aids in computation have been developed which very greatly reduce the amount of labor.[2] Mr. E. R. Wood at the recent meeting of the Educational Research Association presented a remarkable graphical method of obtaining partial correlations which promises to make the powerful method of analysis afforded by partial correlation more widely used for the practical value which it undoubtedly has.

B. R. B.

---

[2] Miner, John R., *Tables of* $\sqrt{1-r^2}$ *and* $1-r^2$. *For use in partial correlation and in trigonometry.* Baltimore: The Johns Hopkins Press, 1922. 49 pp. See also Kelly, Truman Lee. *Tables to facilitate the calculation of partial coefficients and regression equations.* (Bulletin No. 27) Austin, Texas: University of Texas, 1916. 53 pp.

ERRATUM

The reader's attention is called to the article by Walter E. Morgan on Spelling Age in the March number. Toward the end of the article some illustrative examples were given. In each of these, Table II should read Table IV, and Table I should read Table V. It is suggested that our readers get out their March number and make these changes—but only in the three "Examples Illustrating the Computation of the Spelling Age."

Reviews and Abstracts

SMITH, WILLIAM A. *The reading process.* New York: The Macmillan Company, 1922. 267 pp.

For the teacher who desires a better historical background for her work in the teaching of reading, this book will be found a most valuable help. The author traces the development of spoken and written language with special reference to the history of our own English tongue and gives many suggestions regarding the best methods of teaching beginners to read. He shows in a very clear manner the perceptual and interpretative processes in reading, the measurements of which have been determined by certain recent investigations of eye-movements and perceptual span. Several chapters contain descriptive summaries of some recent standard tests and investigations to determine the rates of oral and silent reading, and the comprehension scores. The content of the early readers as well as the more modern ones is discussed.

The purpose of the book, as stated in the preface, is not primarily to give a treatise on methods of teaching reading but rather to give the principles which underlie the method, and to acquaint the teacher with the historical background and linguistic information so necessary for successful teaching. These purposes are most admirably accomplished.

JOSEPH P. O'HERN

Assistant Superintendent, Rochester, N. Y.

PITTMAN, MARVIN S. *Successful teaching in rural schools.* New York: American Book Company, 1922. 294 pp.

One of Mr. Pittman's fundamental beliefs is that teachers are a large agency in their own improvement. He has in "Successful Teaching in Rural Schools" rewritten an account of his earlier study of supervision, making a book unusually desirable for reading circles. While retaining the fundamental outline of his zone plan of supervision, he has elaborated for teacher readers the help he gave in person to the teachers in his supervisory group. "Only enough liberty has been taken with the facts to fill out a fairly general treatment of elementary education as it is applied by the classroom teacher."

There is probably no more readable treatise on elementary education. The book consists of a series of letters from Martha to Hilda. Martha is a rural-school teacher under the supervisory guidance of "Mr. William Hoppes Moore," otherwise Mr. Pittman, and in her letters to her friend, Hilda, she relates in detail the events of the year and their effect upon her own acts and thinking. If Martha is an unusually gifted letter writer, with an uncanny

grasp of the philosophy of education and of life, so much may be pardoned the author. Certainly through Martha's enthusiastic eyes and facile pen he has managed to introduce his readers to practically every important educational idea or practice of the present day,—really to introduce them, not merely to bring reader and idea face to face. Silent reading, standardized tests in school subjects, Binet tests and the I.Q., inductive and deductive teaching, socialized recitations, the project method, democracy and education, health crusaders, improved methods in many school subjects, desirable improvement in teachers' institutes, eugenics—they are all there, sometimes elaborated, sometimes only touched upon, but in every case made meaningful and real through homely comparison or clearly described example.

Hilda is as unusual a reader as Martha is a writer. She meditates upon each letter she receives. She "wonders" the important questions that should be considered in relation to the subject of the letter which she has just read. Questions for study and discussion at the close of a chapter surely were never more disarmingly and attractively presented. "Hilda's Meditations" they are called, and they are followed in each case by a bibliography entitled, "What Hilda Read in Order to Answer Her Questions."

There is no teacher so blind as not to find something in these pages that she will carry over into conduct; and there is probably none so expert that for her they bear no new suggestion or inspiration. The book abounds in concrete accounts of school activities; it abounds too in sound educational philosophy deftly woven into or emerging from the interesting narrative.

FANNIE WYCHE DUNN

Teachers College, Columbia University

NEALE, M. G., and SEVERSON, S. B. *A school building program for the city of Duluth, Minnesota.* Duluth: Board of Education, 1922. 94 pp.

In this report Professor Neale and his associates have used modern scientific technique for determining the future building needs of Duluth. The diagrams and maps help the lay-reader to understand the discussion. The descriptions and the maps showing trends in population are of particular value.

The survey shows: (1) The Duluth school buildings are least efficient in respect to special rooms, fire protection, and toilet systems. (2) Taken as a whole, the school buildings are approximately 62 percent proficient for school purposes as compared with standard school buildings. (3) Playground space in the majority of schools is limited to 50 square feet per pupil enrolled. This is due, in a large degree, to the fact that the peculiar topography of the land in Duluth does not permit the development of adequate play space without large expenditures of money. (4) Seventy-five percent of the school buildings are of the non-fireproof kind, while only 25 percent are fireproof or practically so. The majority of the latter type have been erected since 1910. (5) Inadequate provision has been made in the past for protection against fire and the artificial lighting facilities are only about 75 percent efficient. (6) The majority of the classrooms do not conform to the standard shape and are, as a rule, over-size. (7) Very little, if any, provision has been made in

the schools for special rooms such as auditorium, gymnasium, teachers' rooms, etc. The survey makes valuable suggestions for improving the existing conditions in a general way and also for the improvement of individual buildings.

On the basis of the population studies the population for 1930 is estimated at 122,050 while that for 1940 is estimated at 146,400. With this increase in population the survey recommends certain consolidations in elementary schools now in use so that by 1940 there will be fewer elementary-school buildings in use than are found in Duluth at the present time. The plan for districting recommended, assumes that the distance of pupils from elementary schools will not be too great provided they live within a half-mile radius. The large operating expense of small school plants together with the inadvisability of having a large number of special rooms in such units led the survey committee to recommend this consolidation. In 1940 it is estimated that two senior high schools and seven junior high schools will be necessary to take care of the pupils in grades seven to twelve inclusive.

It was the purpose of the survey to emphasize the mistakes in the location, plan, and construction of school buildings only in so far as they would furnish a basis for scientific planning for the future.

P. R. STEVENSON

Ohio State University

SNEDDEN, DAVID. *Educational sociology.* New York: The Century Company, 1922. 689 pp.

This is a very comprehensive and diversified book of fifty-two chapters. Not only does each chapter contain an appropriate bibliography but there is a special bibliography preceding the index. The whole is divided into four parts: Societies and Social Groups; Social Forces, Processes, and Values; The Sociological Foundations of Education; and The Sociological Foundations of the School Subjects. It is obvious that the first two parts are primarily sociological and the last two primarily educational in nature.

The first three hundred pages are a comprehensive but very sketchy survey of many features of recent sociological literature and discussion. Part I contains short descriptions, numerous definitions, and brief characterizations of recent tendencies affecting a considerable variety of social groups and relationships. One is compelled to survey, even though cursorily, almost every sort of modern problem from that of the local boy gang to that of international imperilism. The book, therefore, resolves itself into a syllabus, a few pages of which outline questions and problems a class may spend many days in discussing.

But this involves the book in the defects of its qualities. It becomes inevitably discursive, piece-meal, disconnected, while its treatment is necessarily inadequate. It is not a book to be read; it is a syllabus which in the hands of a well-trained and widely-read teacher should prove a very valuable guide to a host of interesting if somewhat unrelated problems.

From the standpoint of the sociologist its chief defect lies in its discursiveness and inadequacy of treatment of fundamental matters. The concrete illustrations which the student must draw from his own experience far exceed the theoretical and interpretative viewpoints with which the student is

made acquainted. This is partly a consequence of the author's intent to force the student to draw on his own experiences and partly a result of the extent of territory traversed. One wonders, therefore, how the average class will avoid superficiality of results. From the reviewers point of view there is great difficulty in seeing the woods because of the trees. The book will help classes in education to an analysis of many social groupings and give them an introduction to a variety of sociological concepts, but where will they find the guiding principles that will enable them to put the whole together again so as to give it any significance or meaning?

This last defect appears most obviously if one relate the last two parts of the book to the first two. One searches far in the educational analysis for specific references to any criteria established in the sociological analysis. There is thus an impression of a fatal dichotomy running through the book. Since the book is obviously written for a wide group of public school teachers (the "School Subjects" of Part IV are all below the college) one wonders whether they will be able to bridge this gap between the sociological and the educational with any worth-while generalizations or concepts. Could not the last two parts of the book have been issued separately without appreciable loss of content or viewpoint?

As a syllabus, therefore, the book ought to prove useful in the hands of an adequate instructor. But again one wonders if a book containing less would not have been more serviceable. Why not a book on this great subject devoted to an analysis of the major tendencies of modern society and the significance of these for formal education? Such a book might not be a compendium of encyclopedic diversity; and yet it might say more that was significant in much less space.

F. H. Hankins

Smith College

Wallin, J. E. W. *The achievement of subnormal children in standardized educational tests.* (Miami University Bulletin, XX-7, April, 1922.) Oxford, Ohio: Miami University. 97 pp.

It has often been pointed out that a subnormal child and a normal child of the same mental age are not equal from the achievement standpoint. The normal child has the advantage of developing more rapidly and enriching his knowledge from experiences which may mean little or nothing to the subnormal. On the other hand the subnormal child has the advantage of being older, with the consequence, usually, of having had more drill in fundamentals than the normal child.

In this monograph Dr. Wallin endeavors to discover these differences by means of comparative educational achievement tests in three school subjects—spelling, oral reading, and arithmetic. The tests were made in the special schools of St. Louis in 1918.

Application of the Ayres spelling test (lists A to F and I, L, O) shows increase of spelling ability by grade, mental age, and intelligence classification; although the change is not marked or regular. In some instances the subnormals made greater progress from grade to grade than the normal pupils.

In other cases the normals showed more rapid learning. The Starch Spelling Lists I and II yielded more uniform results, although disturbing factors were also evident. In improvement from grade to grade the subnormals were about equal to the normals. Dr. Wallin concludes that spelling ability above the second-grade level should keep a child out of the classes for the feeble-minded. It is futile, he believes, to try to teach spelling to imbecile children, and it is of doubtful value to morons except of the highest grade. Several cases of word-blindness were found. This condition should be carefully sought lest normal children having this defect be considered subnormal. A positive correlation exists between spelling ability and intelligence. No pronounced sex differences were found.

Gray's Oral Reading Test was applied to 304 pupils. Nearly all children having mental ages of less than five years failed completely. The process of reading was found too hard for these mental levels. Six years is cited as the level below which the teaching of reading to subnormals does not give useful results. Reading ability of third-grade quality or better, however, is sufficient reason for exemption from special classes for the feeble-minded. In reading, as in spelling, cases of special deficiency were found among pupils of greater mental capacity than the achievement test would indicate. For this reason the author cautions against the grading of pupils solely by mental age.

The spiral arithmetic exercises adapted from the Courtis tests gave results similar to those found in reading and spelling. Arithmetic is too difficult for subnormals who have not reached a mental level of six years, although above that level they improve slowly with increasing intelligence.

Dr. Wallin is among the most conservative psychologists in the matter of estimating the proportion of feeble-minded in the schools. The diagnoses, he points out, are too often hastily made, without due regard to special traits, such as word-blindness. He urges the adoption of stricter definitions, and would have the term "feeble-minded" relate to social, not intellectual criteria.

Even by the most conservative classification the proportion of subnormal and feeble-minded children in the public schools is so large that few if any cities are adequately meeting the needs of these pupils. Crime, delinquency, prostitution, and other social evils are very often the consequence of the futile struggle of the incompetent to make a decent living. Through the wider application of eugenic measures the number of such persons may be eventually reduced, but for a long time to come the training and supervision of the feeble-minded will be a necessary function of the public schools. In this monograph Dr. Wallin has contributed some practical knowledge which will be useful in helping the schools adjust these children more effectively to the perplexing problems they will meet in later life.

California Bureau of Juvenile Research J. HAROLD WILLIAMS

REISNER, EDWARD H. *Nationalism and education since 1789*. New York: The Macmillan Company, 1922. 575 pages.

This book is a scholarly, convincing analysis of historical development as related to educational practices in France, Prussia, England, and the United

States. The three great social movements that have conditioned the educational developments of the last century and a half of western history are nationalism, democracy, and the industrial revolution. These three movements, with their many relationships to other aspects of human affairs, are traced in their influences upon social organization and social life. The book is more than a historical account of educational development for it reveals four peoples developing, under the stress and strain of events, an educational system. Each system was designed in turn to perpetuate and develop a nation harmonizing with its ideals. Each of the nations chosen displays a type of social development. Prussia is represented as developing nationalism in education to the extreme; France represents nationalism yielding to the demands of democracy; while England represents a conservative society modifying its educational procedure to meet the pressing needs of the industrial revolution. America is typical of an extreme democracy "gradually responding to the necessity of self-discipline and self-cultivation through public education and only in recent years feeling in marked degree the influence either of the industrial revolution or of nationalism." (p. 3)

The book is continuous exposition and interpretation. There is no cause to plead, no achievement to be praised or blamed. Steadily, with no digression, the author marshals facts and explains whence they came and what results followed. *Compelling* is the only word that describes the author's treatment of his materials. Quotations to show this quality, once started, would be without end. There is, however, one quotation from the "Self-Confessions" of Frederick William III that so aptly expresses the current attitude of some people that it is given in full:

"In respect to the loud and ever louder demand for popular education by means of improved schools, I find myself in a disagreeable position which causes me considerable uneasiness. It must be granted that popular education is the foundation upon which the welfare of the people must rest. A neglected, uncouth, illiterate people can be neither a good nor a happy people. Therefore I have given the good-schools interest a free hand and supported it as far as the economic condition of the state allowed. I have also been pleased to hear the many reports of progress in the Prussian territories. I have also had satisfaction in hearing the comparison made between my own land, in which the great majority of the children receive instruction, and other lands of Europe in which no schools whatever exist.

"But just where educational conditions are most advanced, all kinds of doubts and forebodings force themselves upon me. May one ask himself regarding popular education whether or not it has its limits? If it has no bounds, then we are not justified in interfering with, hindering, or restricting its development, but must let it take its natural course. That, however, I cannot approve without reservations. The answer becomes still more difficult when one wishes to set up limitations and then tries to say where they are to be and whether or not they can be established.

"We do not confer upon the individual or upon society any benefit when we educate him beyond the bounds of his social class and vocation, give him a cultivation which he cannot make use of, and awaken in him pretensions and needs which his lot in life does not allow him to satisfy."

Every person who has any responsibility for developing a view of the function of public education in our country or of determining educational policies should read this book to get the perspective which Dr. Reisner's

treatment gives. Secondary education has been, with us, an upward development of the elementary school and not, as in European countries, "a thing apart from the schools of the people." The unfolding of a flower is shown in our modern motion pictures; education, as a developing social process, is revealed in Dr. Riesner's illuminating pages.

JOHN A. H. KEITH

State Normal School, Indiana, Pennsylvania

The Association of Collegiate Schools of Business. *Social studies in secondary schools.* Chicago: University of Chicago Press, 1922. 117 pp.

In making generally available this little book respecting reforms to be sought in social studies the Commission charged with the task of writing and publishing the report has rendered a distinctive service. Moreover, it has rendered this service not merely to the cause of business education but to the cause of education in general. For certainly if any group of studies in our school curricula needs overhauling, reorganizing, and rebuilding it is the group classed as social studies. This reforming the Commission has fearlessly undertaken to accomplish—or at least to suggest the ways of accomplishing it.

The Commission which has offered the public the latest plan of procedure states frankly it has "no illusions concerning the value of such a study." All it has aimed to do is to give "the Commission's interpretation of the status of the social studies in secondary education and the Commission's hypothesis for their reorganization." In doing this, however, the Commission has been eminently fair in its attitude and wise in its procedure.

In the first place, it has not sought to treat the problem in all its variations, but has concentrated its discussions upon the junior high-school period (seventh, eighth, and ninth grades) and has accepted this unit of school organization as an established fact. In the second place, the Commission makes no claim to having devised a plan that permits of no modifications. In the third place, in making its recommendations, the Commission has kept before it the idea that business education, of whatever kind, is not solely for the sake of personal material aggrandizement but for improved social conditions.

After sketching the history of the efforts heretofore made to secure a modification of the content and methods of the social studies, the Commission points out (1) that "the bulk of the training for business which will be done in this country in our generation will be done by the secondary schools," (2) that statistics show that the typical student in these schools receives less than two and one-half units of instruction in the social studies and that most of this is history, and (3) that the collegiate schools of business, taken as a whole, have not articulated their courses effectively with the work of the secondary schools, and have not provided a well-balanced required program of the social studies.

With these thoughts as background, the Commission declares the social studies ought to be made the backbone of every secondary-school curriculum for the purpose of "opening up the field; of introducing the pupils to society and its problems; of giving the pupils an awareness of what it means to live in organized society, an appreciation of how we do live, and an understanding

of the conditions precedent to living together well." To realize this purpose, the Commission justly urges that instruction in history alone does not suffice, but that sociology, economics, history, government, geography, and current events each has its contributions to make. It therefore proposes that the social studies shall be organized in accordance with the following plan:

Seventh grade: Geography and social science survey—a study of simple industry, simple society, and the transforming influence of scientific knowledge.

Eighth grade: A study of the development and the practices of our own modern social organization.

Ninth grade: A study of the principles of social organization, presented with much "citizenship material."

Stated otherwise the Commission's plan involves: Type studies of social life and interests for the seventh grade, practices and problems of American life for the eighth grade, and principles of sociology for the ninth grade.

The Commission has sought to defend its outline in a few brief paragraphs, and has sketched a good case. Not all will agree that *principles* of social living can be taught effectively to all ninth-grade pupils, but surely the experience in many schools indicates that, under proper conditions, it is possible to teach them to many pupils. Given well-chosen subject material and well-trained teachers there is little doubt but that they can be taught to all. At any rate, to the reviewer, the plan presented by the Commission is the most reasonable one that has yet been published. It is at least worth a trial by all progressive school authorities.

In addition to the body of the report, the volume contains forty-eight pages devoted to a well-selected, well-organized bibliography, designed specifically to guide the further study of busy administrators.

C. O. Davis

University of Michigan

Uhl, W. L. *Scientific determination of the content of the elementary school course in reading.* (University of Wisconsin Studies in the Social Sciences and History) Madison: University of Wisconsin, 1921. 152 pp.

Within the last two years at least a dozen valuable research studies of problems in reading have appeared in monograph form. The study under discussion deals with the content of the elementary-school course in reading. Its purpose is given by the author as follows: "The series of investigations reported in the following chapters was undertaken for the purpose of studying the content of elementary-school reading courses with a view to formulating plans for improving it. This general purpose includes three minor ones, each of which will be considered in detail: first, the elimination of unsuitable reading material; second, the detection of superior reading material; and third, the accurate placement of this superior material."

A questionnaire was sent to a large number of selected teachers in 80 cities located in 25 states in order to secure data for the formulation of a list of the most satisfactory reading selections for each grade. From the returns of this questionnaire a list of about 50 reading selections for each grade was

prepared and certain desirable and undesirable characteristics of reading selections were formulated. A second questionnaire asking the teachers to comment upon the qualities of each selection was sent out. Replies were received from nearly a hundred teachers for each grade in a total of 49 cities. The third part of the study consisted in having 20 specially chosen selections read by 529 pupils from the third to the eighth grade in order to gain information regarding the pupils' likes and dislikes and to determine the relative difficulty of each selection. A comprehension test on each was given for the latter purpose. An additional study was made involving the use of a new type of informational selections taken from the Community and National Life Series.

This elaborate study has yielded some very interesting and valuable conclusions. Many undesirable as well as many highly desirable selections have become established in the reading course. Over-maturity of reading material was mentioned oftener than any other undesirable quality. Interesting action and interesting characters ensure an intrinsic appeal in the selections which portray them. Superior selections usually possess several appeals. Many also interest children of several grades. Poems as well as prose are among the selections which both pupils and teachers designate as superior. Moral qualities are not obtrusive in superior reading matter.

The older informational selections found in some of the readers in current use met with the greatest disfavor. In contrast, the newer informational selections prepared by experts are rated among the most superior selections. The author says, "Interesting informational literature may be written so that it will possess the same important desirable qualities as classical non-informational literature." He also states that it should be written for pupils' use by experts and not culled from masterpieces written for the use of adults.

In the chapter on the "Placement of Reading Selections," the author presents standards for rating reading selections. For each grade, there is given first, a list of desirable qualities with weightings and second, a list of selections scaled as to difficulty and interest. Many selections were found in school readers and courses of study which proved to be undesirable because of over-maturity. Many times these same selections were satisfactory in higher grades.

This study throws a flood of light upon problems involved in the formulation of the elementary-school course in reading. It will be specially valuable to those concerned in curriculum and textbook making and those who have the responsibility of selecting reading books for school use.

C. R. STONE

Gardenville School, St. Louis

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

David W. Griffiths has resigned his position as instructor in mathematics at Carnegie Institute of Technology to become buyer of bonds for a brokerage firm in Minneapolis.

The report of the Survey of Lawrence Township, Mercer County, New Jersey, made by Professors N. L. Engelhardt and E. S. Evenden of Teachers College, Columbia University, has just come from the press of the Division of Field Studies, Institute of Educational Research, Teachers College.

Miss Mary G. Ingersoll, assistant professor of costume economics since 1919, and Miss Sarah J. Weber, instructor in household economics since 1917, have resigned their positions at Margaret Morrison Carnegie College, of Carnegie Institute of Technology, Pittsburgh. Miss Ingersoll has joined the Fisk Teachers Agency, at Berkeley, California, and Miss Weber has gone to continue her studies at Columbia University.

Doctor Fred C. Ayer, director of research, Seattle, Washington, has prepared for the use of the teachers in that school system a bulletin on "The Curriculum." The bulletin includes the results of special investigations under the following headings: curriculum making, time distribution or elementary subjects, scope of curriculum, organization, adaptability to the needs of the children, and extra-curricular activities.

Mr. Charles J. Neagle, Bureau of Measurements, East Stroudsburg (Pa.) State Normal School tested the pupils in his training school with the Illinois General Intelligence Scale and with the Stanford Revision of the Binet Test. He found a very high coefficient of correlation between the results of the two tests, but the number of cases involved were too small to make the results of particular significance.

The Nebraska school statistics for the year 1921-22 show a total expenditure of $29,986,555 of which $15,285,499 was for teachers' salaries. The cost of education per pupil in daily attendance was $119.85. The average daily attendance was 250,785 out of an enrollment of 314,894 and a school population, ages five to twenty inclusive, of 403,146.

The report of the State Department of New Jersey for the past fiscal year shows an operating expenditure for the schools of the state of $42,856,196 of which $30,014,528 was for salaries. The total enrollment of pupils in the public schools, including the evening schools, was 678,734.

Survey Bulletin Volume IV, Number 2, of the Bucyrus (Ohio) city schools continues the type of information which earlier bulletins have given to the teachers and school people of that city. Superintendent Patterson is a thorough convert to the idea of the use of measurement in school supervision. He presents his data in both tabular and graphical form and follows it with comment which should assist the teachers in improving the results of instruction.

Reference was made in these columns some months ago to the Bureau County (Illinois) Standard Test in Special Error Correction. We have recently received from Superintendent Smith a brief statement regarding the results in the use of that test. These results show that Superintendent Smith and his colleagues were quite successful in devising a series of tests equally difficult for the different grades and also one which had diagnostic qualities. Emphasis on language teaching throughout the county this year has been on the basis of the weaknesses revealed by the test.

The annual report of the Carnegie Corporation shows a total expenditure for the past fiscal year of $5,250,000, of which $2,578,000 went to colleges and universities. The Corporation is discontinuing its gifts of libraries, and extending its activities more definitely into the fields of research. The Institute of Economics of Washington, the Food Research Institute of Stanford, and the National Research Council are notable examples of the major interests supported wholly or largely by this corporation.

The report of the educational survey of Indiana made by the General Education Board has recommended the reorganization of the State Department, and the adoption of the county unit system of school government. It proposes better teacher-training facilities and higher requirements; and advocates the equalization of school taxes through increases in state support. The proposed reorganized State Department would include the executive office and the following divisions: schoolhouse planning, teacher training and teacher certification, supervision, statistics and accounting, and compulsory attendance.

Those interested in school-building surveys and programs will find helpful material in Bulletin No. 8 of the Bureau of Research and Education, University of California, entitled "A School-Building Survey and Schoolhousing Program for San Rafael, California," by F. W. Hart, and L. H. Peterson. A careful analysis of the present buildings, progress of children through the schools, growth in population and school enrollment, and the ability of the city to provide adequate housing are the bases upon which the program has been worked out.

The Institute of International Education seems destined to play an important part in education during this generation. Bulletin No. 4 of this Institute is a report on education in China made by Doctor Paul Monroe of Teachers College, Columbia University, who is director of the Far Eastern Bureau of the Institute. The report sets forth briefly but concisely for the benefit of American educational authorities the educational conditions in China today. As we become acquainted through such reports with educational conditions in other countries we shall arrive at a better understanding of conditions in these countries and of our relations thereto.

Mr. P. C. Rogers, Jr., principal of high school, Homer, Louisiana, has recently presented some evidence on the permanence of teaching effort in spelling. At the end of the first semester of the current year he selected a list of fifty words for each grade, distributed equally throughout the words studied during the semester. On this test, his median results in the various grades were either 96 or 98 percent; and the lower quartile in no grade was lower than 88 percent. Perfect papers ran as high as 48 percent in one grade. Apparently the teachers in Mr. Rogers' school are securing a very high quality of work by their teaching effort in this subject.

Volume I, Number 1 of the *Baltimore Bulletin of Education* reached us the latter part of February. This is a twelve-page bulletin full of information and inspiration. While Superintendent West and his entire headquarters staff are listed as the Board of Editors, we note that Doctor Stenquist, director of research, is editor-in-chief. We wish to extend congratulations to Doctor Stenquist and the school people of Baltimore on their enterprise in the publication of this bulletin. If it can reach a large proportion of the thinking citizens of Baltimore we are sure that the Department of Education will be well supported in its activities.

The Pintner-Cunningham Primary Mental Test for kindergarten, first and second grades has just been published by the World Book Company. The test consists of a sixteen-page booklet containing seven tests as follows: common observation, aesthetic differences, associated objects, discrimination of size, picture parts, picture completion, and dot drawing. Tentative mental-age norms cover the range of 48 months to 109 months. The manual of directions contains a key and scoring schedule; but the test manifestly needs a stencil to increase the ease of scoring.

Superintendent Walter R. Siders, Pocatello, Idaho, introduced at the meeting of the Department of Superintendence the idea that "the parent is right." We wonder what value our readers will think this idea has. The business man has learned the lesson that "the customer is always right"; the teacher is too apt to hold the attitude that "the parent is always wrong." "If we could ever adopt in schools the slogan, 'the patron is right,' many of our other problems would be solved. Such an attitude would remove much of

the hostility which deprives schools of moral and even financial support, and causes needless friction."

Doctor Elmer E. Bogart, principal of Morris High School of New York City, recently reported before the New York Society for the Experimental Study of Education concerning the status of Latin in New York State. He affirmed that Latin is taught in all but forty-eight of the 1,038 recognized high schools of the state and that it holds first place among foreign languages on the basis of enrollment. The figures taken May, 1922, report that 93,000 pupils were taking Latin; 79,000, French; 57,000, Spanish; 5,000, German; and 2,000, Greek. Doctor Bogart also showed that the holding power of Latin, as measured by the number of pupils studying it in the third and fourth years, was greater than the holding power of the other foreign languages.

Bulletin Number 49 of the Bureau of Educational Reference and Research, University of Michigan, deals with spelling results on the Ayres-Buckingham Spelling Scale given last October. Results portray a wide variability in achievement of pupils in the different cities of the state with a strong tendency to be below normal expectancy in achievement as designated by the scale. On the basis of these results Doctor Woody has proposed an experiment to throw some light on the permanency of the teaching of spelling. Children in each grade are to be tested upon a special list. This testing is to be followed by the teaching of these words by the teacher's accustomary method. After all the words are taught the children will be retested and again retested after an interval of two months. We hope that a large number of Michigan schools will cooperate in this study since we greatly need extensive studies in the permanency of the teaching results.

Under the direction of Mr. H. T. Manuel, director of educational research, the Colorado State Normal School at Gunnison has announced a scholarship contest open to all high schools in the state. A cup will be given to the school, winning the most points, and medals provided for individual point winners. The two high-school seniors, receiving the greatest number of points, will be given scholarships. The contest will consist of examinations built upon the pattern of standard tests in the high-school subjects of algebra, Latin, American history, civics, typewriting, and correct English. In addition to these there is a general information test involving material from almost all of the common high-school subjects. We wonder if the time is coming when scholastic contests may acquire at least some of the popularity that is now accorded to athletic contests.

A recent circular letter sent to the school superintendents throughout the state by the Connecticut State Board of Education contains the following significant statements: "A recent examination of reports on the amount and kind of professional reading of Connecticut teachers leads to three tentative conclusions: (1) the amount of reading and study voluntarily assumed is

small, the average being less than one book a year; (2) the educational periodical literature is of the device-suggesting type, containing too many unauthentic suggestions often based upon an unsound psychology of teaching; and (3) worth-while books and periodical magazine articles, e. g., those discussing principles, special methods, relation between courses of study, provisions for general curriculum, aims, etc., are too rarely matters of voluntary selection by teachers." It is highly probable that a similar study of teachers' reading in other states would show a like condition.

The N. E. A. Commission on the Coordination of Research Agencies has a sub-committee on the indexing and classification of educational writings. Professor Thomas H. Briggs is chairman of this sub-committee. During the past few months the work of the sub-committee has been carried on in part at the Bureau of Educational Research at Ohio State University. Miss Margaret Doherty, reference assistant in the Bureau, has been trying out certain classification lists and is now working on a list of subject heads to be used in indexing educational material. This list will be issued in mimeographed or printed form for the criticisms and suggestions of educational people. Based on their responses a revised list will be issued for the use of libraries, bureaus, and offices which handle educational writings. This work is regarded as important because it will standardize the indexing of books, bulletins, pamphlets, reports, and periodicals, and will permit the interchange of bibliographies and citations in ways that will be understood by all users of the subject list.

A preliminary notice has been received of the Second International Conference of Education to be held at Institution Des Essarts, Territet, Switzerland, August 2-15, 1923. The aim of the promoters of the conference is to discover and disseminate the most ideal methods of education, believing that by so doing they will not only be helping the child to develop its own capacities to the higher degree of efficiency but also that they will be employing the only possible means of securing the best type of future citizen. The following representatives of the different countries are on the conference program:

Austria	Dr. Otto Rommel	Bundeserziehungstalten
	Professor Cizek	Arts and Crafts
	Frau Scheu-Riesz	World Literature for Children
Belgium	Dr. Decroly	Intellectual
England	Mr. Cloudesley Brereton	
	Mr. Henry Wilson	Arts and Crafts
	Mrs. Beatrice Ensor	Schools of Tomorrow
France	M. Bertier	Ecoles Nouvelles
	Mme. Jouenne	Open-air Schools
Germany	Dr. Elizabeth Rotten	The New Education
Italy	Dr. Maria Montessori	The Montessori Method
Switzerland	Dr. Jung	Analytical Psychology
	Dr. Ferrière	Ecoles Nouvelles
	M. Jaques-Dalcroze	Eurhythmics
	M. Hermann Tobler	Landerziehungsheime

Superintendent William E. Stark of Hackensack, New Jersey, makes some valuable distinctions in connection with the duties of the superintendent. He rightly points out that one of the most fruitful sources of improvement in education lies in the latent ability of the teaching staff. It is the superintendent's duty to develop this ability. He cannot do this by the "six-horse team theory" in virtue of which he "holds the reins and plies the whip." He cannot do it by the "Pollyanna theory" by wearing a permanent smile and commending everything. Moreover, the "government inspection theory" will not do. This theory makes the superintendent a plain-clothes man looking for infractions of the law and applying the government label. According to Superintendent Stark the superintendent gets the best from his teachers when he becomes a professional leader. He cannot inspect everything. He must have confidence in his teachers and give them responsibility. He must train his teachers to recognize and work out their problems for themselves. Under such circumstances, he will be neither a driver, a nurse, nor an inspector. He will be captain of a team.

We lately received an interesting volume from Superintendent Milton C. Potter of Milwaukee, entitled "Projects and Games in the Primary Grades." It is a reprint in revised and illustrated form of a bulletin issued, as we understand, in 1921. It is prepared by the primary teachers of the Milwaukee public schools under the direction of Margaret Canty, assistant superintendent in charge of primary instruction. The first part consists of projects applicable to the first three grades and definitely related to the prescribed curriculum. For example, in the first grade a few of the projects are *a mimic theater, the doll project, a birthday party, the picnic,* and *Easter in a primary room.* Miss Canty contributes a synopsis of the project method and a plan for project teaching in the first grade. Dr. W. W. Theisen has a few paragraphs on "Standards for Judging School Projects."

The second part of the book describes a large number of games to be used in reading, language, arithmetic, spelling, and geography classes. A bibliography on projects and games is appended.

An interesting study of the elementary-school curriculum in America previous to 1800 has been made by R. D. Russell, a graduate student in education at the University of Iowa. The outstanding facts are as follows: "In the teaching of reading the Hornbook, Battledoor, Primer, Catechism, Psalter, Testament, and Bible were used. At the present time only three copies of American hornbooks are known to be in existence. The names of the primers used in America before the New England Primer was published are not known; but it is believed that Reverend John Eliot composed one in about 1668. The first reader, in a modern sense of the word, was by Noah Webster in 1783. Possibly the first spelling book used in Colonial America was the 'English School Master' written by Edmund Coote and published in 1596. The contents of this book are as follows: 32 pages given to alphabet and spelling; 18 pages to a short catechism, prayers, and psalms; 5 pages to

chronology; 2 to writing copies; 2 to arithmetic; and a few pages to a list of especially hard words alphabetically arranged.

"The first arithmetic printed in English was written by Robert Record and published in 1540; and the first one by an American author was by Isaac Greenwood, published in 1729. The first geography by an American author was 'Georgraphy Made Easy' by Morse in 1784."

In a recent circular to teachers sent out by the Department of Public Instruction of Hawaii, seven basic principles concerning the school system are laid down. Briefly, they are as follows: (1) The primary aim of our public school system is to train for effective American citizenship. This training includes the mastery of the "three R's," a friendly attitude toward industry and American home-making, and a working knowledge of American history, civics, and the ideals of democracy. (2) The American public school system endeavors to provide (for those who are qualified), an unbroken pathway from kindergarten to University. (3) American homes can be best made and maintained by women trained in American home-making. (4) American homes are dependent on dignifying industry . . . . This means industrial and agricultural education, trade schools, part-time classes, continuation classes, and vocational guidance. (5) Physical education is of strategic importance as the productiveness and happiness of a community is determined in large measure by its health. (6) Wholesome moral and ethical teaching should pervade the curriculum, as the final products of education are character and personality. (7) The cost of education is chargeable, not alone to the individual or child, but to the State, which maintains the public schools in part to protect its own future, and to preserve its social institutions against ignorance and all the forces of evil.

A preliminary report on the mental capacity of Japanese children in California by M. L. Darsie, assistant director of Stanford University Research, presents in concise form certain interesting data. The general conclusion reached is that a large handicap is very evident in most of the test performances and its effect on final scores is very hard to estimate. It is fair to assume, however, that it probably tends somewhat to lower the I. Q. Taking all the results into account, the following conclusions seem fully justified:

1. Japanese in California are, as a group, somewhat inferior in intelligence to Northern Europeans but superior to Southern Europeans.
2. In application and capacity to learn they are probably superior to any European race in America, as well as superior to native Americans.
3. In social-moral traits they are fully equal, and in many respects probably superior to the average child of other races in California, as judged by the teachers, this being true with respect to native American children as well.
4. On the basis of teachers' ratings, Japanese children seem to be very superior in school deportment; slightly superior in application, appreciation of beauty, permanence of moods, desire to excel, freedom from vanity and conscientiousness, as well as the school subjects of drawing and painting, music, penmanship, spelling, arithmetic, and physical training.

"Sound-Proof Partitions" by F. R. Watson, published as Bulletin No. 127 of the Engineering Experiment Station, University of Illinois, contains a report on an elaborate investigation of the acoustic properties of various building materials, with practical applications. Two or three items in the general conclusions are of interest to school people.

"Sound-proof partitions should be as rigid and free from air passages as possible. For effective sound-proofing of a group of rooms, the partitions, floors, and ceilings between adjacent rooms should be made continuous and rigid. Any necessary openings for pipes, ventilators, doors, and windows should be placed in outside or corridor walls where a leakage of sound will be less objectionable.

"The vibrations of a motor, fastened to a floor, may be minimized by placing a layer of hairfelt, or similar air-filled material between the supporting base and the floor. The edges of the floor should be insulated from the walls by felt or similar material. Without some efficient control of the transference of sound through the ventilation system it is a waste of effort to construct sound-proof walls, double doors, and other contrivances for insulation."

Some attention to such information as the above in the construction of many of our school buildings would have prevented the noisy conditions which are now prevalent.

We lately received from Miss Velda Bamesberger, director of educational statistics of Okmulgee, Oklahoma, an unusual cost analysis of the high-school teaching in that city. Unit costs were not only computed on the basis of attendance but also on the membership and enrollment. The school accounting people—tending as they do to have the accountant's point of view—quite generally advocate basing unit costs upon attendance. They say that attendance reports the "solid service" of the school. Every school man, however, knows that there is a type of service of undoubted solidity which is rendered with reference to all the pupils who are brought to the school. The fact of their absence does not necessarily diminish this service. Indeed, it may cause the exertions of the teachers to be even greater than they would be if all the pupils were present every day.

Miss Bamesberger makes a searching analysis not only with reference to salary costs but also with reference to other instructional items. Her recommendations regarding sizes of classes are particularly good. After analyzing her figures she combines them and finds that the average annual cost per high-school course is $19.08. On the assumption that a child takes four courses, this would amount to $76.32 as the estimated expense of teaching a pupil for a year. To this, however, should be added the cost of work in physical education and music, and the expense of general maintenance. This would run the cost up to considerably over a hundred dollars per pupil. The cost per elementary-school pupil amounts to about $60.

Educational Research Association

(FRANK N. FREEMAN, *Secretary and Editor*)

Doctor W. J. Osburn, director of Educational Measurements, Wisconsin State Department of Public Instruction, has extended his practical remedial investigation to include a diagnostic study of arithmetic errors. In his mimeographed report "Diagnostic and Remedial Treatment for Errors in Arithmetic Reasoning" he says: "Progress in the study of errors has been long retarded by the idea that there is the utmost diversity in the mistakes of children. The actual evidence is far to the contrary. The same wrong answers occur in all types of schools and over wide areas to such an extent that it is possible to forecast how many pupils in any individual school will make certain mistakes."

A remedial program according to Mr. Osburn must include: (1) a list of typical problems which have been widely used and on which lists of wrong answers are available; (2) an identification and classification of the most serious disabilities; (3) a list of the wrong answers which are symptomatic of each disability in each type of problem; and (4) the remedy to apply after the specific disabilities are known and identified.

In his investigation Doctor Osburn used 30 problems from the Buckingham Problem Scale in Arithmetic. Based upon a study of 30,000 errors made by 6,000 children in eighteen counties and one large city, he gives the approximate percent of frequency for the following errors:

Types of Errors	Approximate Frequency Percent	
(1) Total failure to comprehend the problem	30	
(2) Procedure partly correct but with the omission of one or two essential elements (This type will be designated as partiol response hereafter)	20	
(3) Ignorance of fundamental quantitative relations	10	
Total		60
(4) Errors in fundamentals	20	
(5) Miscellaneous errors	2	
(6) Errors whose cause could not be discovered	18	
Total		40

Mr. Osburn takes up each of the 30 problems which he used and gives for each the typically wrong answers and the disabilities of which they are the symptoms. He discusses general symptoms and remedial instruction. He also includes several silent reading exercises to test the pupil's ability to read arithmetic problems before attempting a solution. Mr. Osburn promises two further circulars on remedial work in arithmetic which are now in preparation.

The Westchester County Group Testing Program

Some months ago the people of Westchester County, through an educational conference, organized a Committee of One Hundred to study the educa-

tional problems of the county. This committee organized a number of sub-committees, one of which was known as the sub-committee on Special Schools and Classes. Mrs. Caspar Whitney was made chairman. After a number of meetings and the collection of considerable data, the sub-committee decided to determine through the use of intelligence tests, what need there is in the county for the organization of special schools and classes. As a first step a policy committee consisting of two city superintendents, a district superintendent, a high-school principal, and a psychologist was appointed to devise a special program. They proposed that all pupils in grades III to VI inclusive of the entire county should be tested with a group intelligence test and secondly, that all pupils who appeared subnormal according to the group test should later be given an individual psychological examination.

As a result of the policy committee's recommendation an executive committee was appointed consisting of Doctor E. M. Quittmeyer, principal Drum Hill High School, Peekskill; Mr. L. F. Dodge, superintendent of Schools, Yonkers; and Doctor Elizabeth I. Adamson, psychologist, White Plains. The executive committee asked for the cooperation of the State Education Department. Arrangements were made for Doctor J. Cayce Morrison, specialist in educational measurements, to direct the group testing program and for the individual testing to be done by a staff of examiners under the direction of Doctor W. B. Cornell, mental diagnostician of the State Education Department.

The group testing program was definitely organized at a meeting with principals, superintendents, and examiners at the White Plains High School, Tuesday, November 14. On the two days following, the director of the program met with groups of examiners in Yonkers and White Plains for further demonstration and discussion. Later, similar demonstration conferences were held in New Rochelle and Port Chester. The director selected as assistant examiners people who had had previous experience in giving group tests and who had had college or university courses in educational measurements. In nearly every city or village of Westchester County it was found that one or more members of the teaching or supervisory staff had had such experience or training. Following the demonstration conferences the director of the program prepared a summary of the discussion in the form of bulletins on organization and directions to examiners. It was agreed that classroom teachers should rate the papers of their own pupils under the general direction of the principal of the school and should make their class record sheets in duplicate. Later the papers and record sheets will be reviewed by a special staff selected by the director of the program, and reports will be prepared and submitted to each superintendent concerned and to the Executive Committee.

It is estimated that this group testing program will include approximately 25,000 children. The Westchester County superintendents and principals are also considering the use that they can make of the test data for pupils who are found to be normal or above normal in ability to do school work. In view of the contemplated measures, this Westchester County Group Testing Program promises to be the most important program of mental testing yet undertaken in New York State.

J. CAYCE MORRISON

State Education Department, Albany, New York